

CERGE
Center for Economics Research and Graduate Education
Charles University Prague



Essays in Behavioral and Development Economics

Vojtěch Bartoš

Dissertation

Prague, August 2016

Vojtěch Bartoš

Essays in Behavioral and Development Economics

Dissertation

Prague, August 2016

Dissertation Committee

MICHAL BAUER (CERGE–EI; chair)

RANDALL FILER (Hunter College, CUNY)

PETER KATUŠČÁK (RWTH Aachen)

KARNA BASU (Hunter College, CUNY)

Referees

PETER MARTINSSON (University of Gothenburg)

JONATHAN MORDUCH (NYU Wagner Graduate School of Public Service, New York University)

Table of Contents

Acknowledgments	ix
Introduction	1
1 Seasonal Scarcity and Sharing Norms	7
1.1 Introduction	8
1.2 Related Literature	12
1.3 Experimental Design	14
1.3.1 Sample Selection	14
1.3.2 Seasonal Effects	15
1.3.3 Experimental Tasks	16
1.3.4 Experimental Procedures	17
1.4 Results	19
1.4.1 Temporal Stability of Sharing Behavior	19
1.4.2 Temporal Stability of Norm Enforcement	22
1.5 Discussion	25
1.5.1 Determinants of Seasonal Changes in Norm Enforcement	25
1.5.2 Individual Preferences or Social Norms as Determinants of Punish- ment Behavior	29
1.5.3 Generalizability	30
1.5.4 Potential Confounds	31
1.6 Concluding Remarks	32
1.A Appendix 1	46
1.A.1 Image Documentation	57
1.A.2 Experiment Instructions	58
2 Contract Enforcement and Trustworthiness Across Ethnic Groups: Ex- perimental Evidence from Northern Afghanistan	61
2.1 Introduction	62
2.2 Experimental Design	66
2.2.1 Experimental Games	66
2.2.2 Treatments	68
2.2.3 Procedure	68

2.3	Theoretical Background	71
2.4	Results	73
2.4.1	Trustee Experimental Results	73
2.4.2	The Behavioral Effect of Sanctions	76
2.4.3	Sanctioning and Fairness	78
2.4.4	Investor Experimental Results	80
2.5	Discussion	83
2.A	Appendix 2	93
2.A.1	Experiment Instructions	110
3	Attention Discrimination:	
	Theory and Field Experiments with Monitoring Information Acquisition	123
3.1	Introduction	124
3.2	The Model of Attention Discrimination	129
3.2.1	Set-up of the Model	129
3.2.2	Effects of Preferences and Beliefs on Attention Allocation	131
3.2.3	Endogenous Attention and Discrimination in Selection Decisions	134
3.3	Field Experiment in the Rental Housing Market	137
3.3.1	Experimental Design	138
3.3.2	Sample Selection and Data	141
3.3.3	Results	142
3.4	Field Experiment in the Labor Market—Czech Republic	148
3.4.1	Experimental Design	148
3.4.2	Sample Selection and Data	149
3.4.3	Results	150
3.5	Field Experiment in the Labor Market—Germany	155
3.5.1	Experimental Design	155
3.5.2	Results	157
3.6	Links to Theories	158
3.7	Concluding Remarks	161
3.A	Appendix 3	171
3.A.1	Supplementary Material to Section 3.2	171
3.A.2	Supplementary Material to Sections 3.3-3.5	172
	Bibliography	191

Abstract

In the first chapter, I examine the effect of scarcity on sharing norms and preferences. Sharing provides one of few sources of insurance in poor communities. It gains prominence during adverse shocks, often largely aggregate, when it is also costliest for individuals to share. Yet, how scarcity affects individual willingness to share and willingness to enforce sharing from others, an important ingredient in sustaining prosocial behavior, is little understood. This is what this paper examines. I conduct repeated within-subject lab-in-the-field experiments among Afghan subsistence farmers during a lean and a post-harvest season of relative plenty. These farmers experience seasonal scarcities annually. Using dictator and third party punishment games I separate individual sharing behavior from the enforcement of sharing norms. While sharing exhibits a high degree of temporal stability at both the aggregate, and, to a large extent, the individual level, enforcement of sharing norms is substantially weaker during the lean season. The findings suggest that farmers are capable of sustaining mutual sharing through transitory periods of scarcity. It remains an open question whether exposure to unexpected shocks or prolonged periods of scarcity might result in the breakdown of prosociality due to loosened sharing norms enforcement on a community level.

In the second chapter, we study how the availability and use of a specific formal institution— a financial sanction – affects trust, trustworthiness, and moral intentions towards co-ethnics and non-co-ethnics using an economic experiment involving 420 adult males from peri-urban areas in Afghanistan. In contrast to previous studies on the behavioral effects of financial incentives, our subjects have little experience with formal institutions. We use a trust game with a requested back-transfer in which the investor can choose to impose a financial sanction for non-compliance. The sanction is costly to the trustee but cost-less to the investor. While sanctioning increases back-transfers in cross-ethnic pairs, it does not in co-ethnic pairs. Our results suggest that financial sanctions may crowd out moral incentives more strongly among one’s own group, but have a much smaller behavioral effect when applied to individuals from a different ethnic group. The results have important implications for understanding how formal institutions affect cooperation in ethnically heterogeneous settings.

In the third chapter, we integrate tools to monitor information acquisition in field experiments on discrimination and examine whether gaps already arise when decision-makers choose an effort level for reading an application. In both of the countries we study, negatively stereotyped minority names reduce employers’ effort to inspecting re-

sumes. In contrast, minority names increase information acquisition in the rental housing market. Both results are consistent with a model of endogenous allocation of costly attention, which magnifies the role of prior beliefs and preferences beyond that considered in standard models of discrimination. The findings have implications for the magnitude of discrimination, returns to human capital and policy.

Abstrakt

V první kapitole zkoumám, jaký efekt má nedostatek na normy a preference, podle kterých se lidé dělí s ostatními. Sdílení zdrojů je jedním z mála způsobů, jakým se lidé v chudých komunitách zajišťují. Nejvýznamnějším se stává během negativních šoků, většinou agregátních, kdy je však sdílení pro jednotlivce nejnákladnější. Zatím se ví jen málo o tom, jak nedostatek ovlivňuje ochotu jednotlivců sdílet a ochotu vynucovat sdílení od ostatních, což je zásadní podmínkou pro zachování prosociálního chování. Tato kapitola se věnuje právě tématu sdílení na základě „lab-in-the-field“ experimentů, kterých se účastnili farmáři v chudých venkovských oblastech severního Afghánistánu. Do experimentů se zapojili opakovaně – v období hladu (jež zažívají každoročně) a v období relativního dostatku po sklizni. Pomocí „hry na diktátora“ (dictator game) a hry s „nestranným trestajícím“ (third-party punishment) je v experimentech rozlišována individuální ochota sdílet od ochoty vynucovat sdílení od ostatních. Zatímco sdílení vykazuje značnou míru stability v čase jak na agregátní, tak na individuální úrovni, vynucování sdílení je v období hladu značně oslabené. Toto zjištění ukazuje, že inkriminovaní farmáři jsou schopni udržovat vzájemné sdílení během přechodných období nedostatku. Zůstává však otázkou, zda u komunit, jež jsou vystaveny nečekaným šokům či déle trvajícím obdobím nedostatku, může dojít kvůli oslabenému vynucování norem sdílení k rozpadu prosociálního chování.

Ve druhé kapitole zkoumáme pomocí ekonomických experimentů s 420-ti dospělými muži z příměstských oblastí v Afghánistánu, jak dostupnost a použití specifické formální instituce – finanční sankce – ovlivňuje důvěru, důvěryhodnost a morální úmysly vůči jednotlivcům z vlastní či cizí etnické skupiny. Narozdíl od předchozích studií zkoumajících behaviorální aspekty finančních pobídek, účastníci našich experimentů mají jen málo zkušeností s formálními institucemi. Používáme „hru na důvěru“ (trust game) s požadovaným zpětným transferem, ve které si investor může zvolit uvalení finanční sankce při nesplnění podmínek. Sankce je nákladná pro „zplnomocněnce“ (trustee), ale je zdarma pro investora. Zatímco použití sankce zvyšuje zpětné transfery u dvojic z různých etnických skupin, transfery u dvojic ze stejné etnické skupiny zůstávají neměnné. Naše výsledky naznačují, že finanční sankce mohou vytěsnit morální úmysly při jednání s jednotlivci z vlastní etnické skupiny, ale mají mnohem slabší behaviorální efekt při jednání s jednotlivci z jiné etnické skupiny. Tyto výsledky prohlubují naše chápání dopadů formálních institucí na spolupráci v etnicky heterogenním prostředí.

Ve třetí kapitole využíváme v internetovém experimentu nástroje umožňující moni-

torovat, jak jsou zpracovávány informace při posuzování žádostí o práci nebo o pronájem. Zkoumáme, kdy vznikají rozdíly v tom, jakou odpověď uchazeč od člověka rozhodujícího o jeho přijetí nebo nepřijetí dostane – zda rozdíly nastávají už tehdy, když člověk zodpovědný za rozhodnutí zvažuje, kolik úsilí věnuje čtení dané žádosti. V obou zemích, které studujeme, snižují jména typická pro jednotlivce z negativně stereotypizovaných skupin úsilí zaměstnavatelů zkoumat jejich životopisy. Naopak jména, která evokují příslušnost k menšině, zvyšují úsilí při vyhledávání informací na trhu s nájemním bydlením. Oba výsledky jsou konzistentní s modelem endogenní alokace nákladné pozornosti, což zesiluje význam apriorních přesvědčení a preferencí nad rámec těch uvažovaných standardními modely diskriminace. Naše zjištění mají možné dopady pro míru diskriminace, návratnost vzdělání a politická opatření.

Acknowledgments

A large number of people contributed to this dissertation in some way and I would like to express my gratitude to all of them.

First, I would like to thank my supervisor, Michal Bauer. He got me interested in the study of behavioral and development economics. His inspiration and constant encouragement also motivated me to pursue my doctoral studies and kept me going to the finish line resulting in this volume.

Second, I am greatly indebted to the members of my dissertation committee: Peter Katuščák, Randall K. Filer, and Karna Basu. They have devoted lots of their valuable time to me. Our many discussions have helped to improve my work and broaden my interest and knowledge in economics.

I am very much indebted to my great officemates at CERGE-EI (and at the Charles University): Jana Cahlíková, Klára Kalíšková, Dáša Celik-Katreniaková, Lubomír Cingl, and Ian Levely, who kept me motivated and also contributed greatly to my work. The endless academic discussions—be it in the office, in numerous bars, or in the middle of the mountains after a long day of climbing—kept me going. I am delighted to see that we are able to sustain these discussions even though we are finishing our studies and moving on with our lives and work..

This dissertation would not be complete without my co-authors: I have already thanked Michal Bauer and Ian Levely, but but I also gratefully acknowledge Julie Chytilová and Filip Matějka.

I partially wrote this dissertation during my research stay at New York University in the Fall of 2014. Jonathan Morduch was extremely kind to invite me to New York and he was an excellent host. Discussions with him were very insightful. My stay also provided me with multiple opportunities to present my work at various seminars and to discuss my research with faculty members at NYU, Columbia University, CUNY, Princeton University, and at Rutgers University.

I benefited from numerous discussions with many other excellent academics throughout my studies. I am especially grateful to Abigail Barr, Alexander W. Cappelen, Subhasish M. Chowdhury, Lee Cronk, Dirk Engelmann, Ernst Fehr, Guillaume Frechette, Ira Gang, Stěpán Jurajda, Andreas Ortmann, Gérard Roland, Ondřej Rydval, Jakub Steiner, and Bertil Tungodden. Audiences at numerous universities, workshops, seminars, and intensive courses have also contributed to the work.

Experimental research also requires assistance with conducting experiments and logistics. I was lucky to receive excellent research assistance from Kateřina Boušková, Ahmad Qais Daneshjo, Hadia Essazada, Lydia Hähnel, Vít Hradil, Mohibullah Mutahed, Iva Pejsarová, Yar Mohammad Rajabi, Akram Rasaa, Lenka Švejdová, Kamran Shahzad, and Viktor Zeisel. I am deeply indebted to the organization People in Need for providing me with logistical support and security during my research stays in Afghanistan.

My writing would not be as polished without the great effort and patience of the English department at CERGE-EI. Namely, I would like to thank Andrea Downing, Robin-Eliece Mercury, and Deborah Nováková for all they have shown and taught me.

Others will surely forgive, but I must thank most of all my wonderful wife Kristýna who accompanied me through all the ups and downs, my lovely son Šimon who allowed me to finish my writing and smiled at me on the right occasions, and all my family for their unconditional support, love and patience.

Thank you! Děkuji! Ďakujem! Danke! Merci beaucoup! Takk! Tashakor! Dhanyavaad!

Financial support by the Grant Agency of Charles University (grants no. 71610 and no. 46813), the Global Development Network (grants no. RRC10/105 and no. RRC13/11), and the Czech Science Foundation (grant no. 13-20217S) is gratefully acknowledged.

All errors remaining in this text are the responsibility of the author.

Czech Republic, Prague
August 2016

Vojtěch Bartoš

Introduction

This dissertation contributes to the rapidly expanding literature linking behavioral economics with the field of development economics. In particular, it addresses several channels through which certain aspects of human behavior contribute to the persistence of poverty, underdevelopment, and inequality. My coauthors and I examine the role of social preferences, and the role of limited attention to information as underlying factors.

The first two chapters aim to extend the literature on state-dependent social preferences, and its role in poverty and economic underdevelopment. The first chapter examines the response of sharing preferences and norms enforcement to resource scarcity. Since informal sharing provides an important social safety net in poor village communities, understanding the responses of its fundamental determinants to shocks is of great importance. The second chapter examines the response of trustworthiness, a crucial component in imperfect contractual agreements, to the introduction of a specific type of formal institution. We artificially construct ethnically homogenous and heterogenous groups to better understand the role of group composition in such behavioral responses. The third chapter studies the role of group-specific acquisition of information about applicants in different types of markets and documents its role in magnifying the discrimination of ethnic minorities beyond the predictions of traditional models of discrimination, further enhancing social inequality.

Methodologically, all chapters use economic experiments to test the issues outlined above. The first two chapters are *artefactual field experiments*¹ conducted with partici-

¹See Harrison and List (2004) for the taxonomy of field experiments.

pants in rural and urban Afghanistan, respectively. Although the participants make their decisions in a controlled laboratory setting, they are expected to bring their experience from their lives that are relevant to the issues studied. The first chapter uses tools from experimental economics to measure individual behavior that would otherwise be difficult to measure and compare across two different states—that of poverty and that of relative plenty—using observational data. The same farmers in the sample are exposed to both states at different points of time, which allows for causal inference. In contrast, the second chapter is a proper experiment with a control and treatment group assigned in a pop-up field laboratory. The third chapter takes one further step towards the field and can be classified as a collection of *natural field experiments*, this is that the participants made decisions in their natural environments—the actual Czech and German HR managers and landlords responded to email applications over the Internet—and are not aware that they are part of an experiment when making decisions. When studying sensitive issues, such as discrimination, it is particularly important that the participants are unaware of being observed. Otherwise they would be unlikely to reveal their true preferences. This chapter also brings a theoretical contribution: the experiments are designed to test the theory of *attention discrimination*, a novel concept we propose.

In what follows I describe the particular chapters in more detail, starting with the two chapters contributing to the literature on social preferences and finishing with the chapter on the role of attention to information in discrimination.

While the presence of social preferences² helps establish and sustain cooperation (Fehr and Gächter 2002), decreases the costs of economic transactions (Arrow 1972; Tirole 2011), affects redistribution policies (Fisman, Jakiela, and Kariv 2015; Fisman et al. 2015), or influences labor-market relations (Fehr, Kirchsteiger, and Riedl 1993), all contributing to more equitable (and often richer) societies, little is understood as to how social preferences and social norms are shaped by poverty. The aim of the first chapter is to fill this gap. It is the first study to examine the effect of poverty on sharing preferences and on the enforcement of sharing norms. Informal sharing is an important coping mechanism in poor communities where formal insurance is virtually non-existent. The enforcement of sharing norms has previously been identified as central for sustaining prosocial behavior in groups (Gintis 2000; Henrich et al. 2006). Yet there is a trade-off:

²It is now well established that incorporating others' utilities into one's own helps explain some well documented phenomena that contradict or cannot be fully explained by the self-regarding models, typically assumed in neo-classical economics.

the need for sharing and its enforcement is highest at the time when such actions are individually most expensive. The question I ask gains importance as recent research has established that poverty fundamentally changes individual behavior. Poverty reduces individual aspirations and leads to underestimation of returns to education (Jensen 2010), lowers cognitive ability (Mani et al. 2013), increases stress and may thus reduce willingness to take risky decisions (Haushofer and Fehr 2014), and increases anti-social behavior (Prediger, Vollan, and Herrmann 2014).

I exploit the fact that farmers in rural Afghanistan are affected by seasonal scarcities every year—as are millions in many other developing countries—before they collect their harvest from which they have to live throughout the year. To measure the stability of sharing preferences and of enforcement of sharing norms, I follow the same farmers across seasons and repeatedly administer a dictator game and a dictator game with a third party punishment opportunity. I find that individual sharing preferences are temporarily stable, but the willingness to engage in third-party punishment drops substantially in the period of scarcity. Although the farmers sustain mutual sharing during the observed transitory periods of scarcity, it remains an open question whether exposure to unexpected shocks or prolonged periods of scarcity might result in the breakdown of prosociality due to loosened sharing norms on a community level.

The second chapter (a joint work with Ian Levely) examines social preferences from a different perspective. It builds on two streams of literature: one that calls attention to the crucial role of trust in contractual agreements (Arrow 1972; Tirole 2011), and the second showing that trust and cooperation are lower in ethnically heterogeneous societies (Alesina and La Ferrara 2005). Miguel and Gugerty (2005) link this finding to the lower ability of ethnically heterogeneous communities to engage in collective action. This generates a puzzling conjecture: the emergence of formal institutions complementing trust-based interactions might be more difficult in ethnically heterogeneous societies, which would benefit from them the most. This situation might prove to be one of the contributing factors to the chronic underdevelopment of ethnically heterogeneous societies.

In particular, we study how the availability of a formal sanctioning mechanism in an otherwise trust-based interaction affects trust, trustworthiness, and moral intentions towards co-ethnics and non-co-ethnics in peri-urban areas in Afghanistan. We find that the crowding-out of intrinsic motivations, by use of a formal sanction, is stronger when the interaction involves individuals from the same ethnic group. This implies that introduction of formal contracts might be more efficient when implemented between ethnic

groups rather than within existing co-ethnic communities, even though the ethnically heterogeneous communities might struggle to establish such institutions on their own in the first place. Interestingly, compared to results from developed societies with established formal institutions (e.g., Fehr and Rockenbach 2003; Fehr and List 2004; Houser et al. 2008), the use of sanctions in our study triggers a much milder detrimental effect on intrinsic motivations even in the co-ethnic interactions.

The role of ethnicity in interpersonal behavior ties together the second and the third chapter. While the second chapter is interested in the role of ethnicity on behavioral responses to the introduction of a formal institution, the third chapter examines how ethnic minorities are discriminated against by majority groups. Whereas the first models of discrimination attribute discriminatory behavior to the role of individual distaste towards minority groups (Becker 1971), other theories consider lack of information about specific individuals as the main contributing factor. This is the main premise of statistical discrimination models (Phelps 1972; Arrow 1973). These models assume that limited information induces individuals to extract information about others from easily accessible sources, such as from their group attributes. Such information processing makes candidates from groups with worse average quality less desirable. Any information about a specific individual, however, can improve the accuracy in assessing the true quality of that individual beyond the quality expected based on the a priori best available signal: the average quality within a group. The model, however, operates with an underlying assumption that any individual-specific information is processed completely.

In the third chapter (a joint work with Michal Bauer, Julie Chytilová, and Filip Matějka), we relax the assumption of full information acquisition and we propose a concept of *attention discrimination*. This arises due to decision makers' endogenous allocation of costly attention to information provided by applicants from different groups. We test the model by integrating tools monitoring information acquisition in field experiments on discrimination in the Czech Republic and in Germany. To our knowledge, this is the first Internet field experiment to use such tools in field experiments. In the labor market we find that negatively stereotyped minority names reduce employers' effort to inspect applicants' resumes, while we observe increased information acquisition about prospective minority-named tenants by landlords in the rental housing market. These results are consistent with our model and imply that prior beliefs about an individual based on her group characteristics play an even greater role than standard models of discrimination (Becker 1971; Phelps 1972) would predict. The findings have implications

for policy and suggest that postponing the revelation of the group attribute—for example by introducing blind resumes—may reduce the attention asymmetry across groups.

A struggle to understand the sources of poverty, underdevelopment and inequality, and to find ways to improve the situation of the most unfortunate people have been the main motivating forces for me throughout my studies of economics. This dissertation offers several insights in this direction. It also opens further academic questions for exploration, and it suggests some policy advice that might deliver more immediate results.

Chapter 1

Seasonal Scarcity and Sharing Norms

Vojtěch Bartoš¹

Abstract

Sharing provides one of the few sources of insurance in poor communities. It gains importance during adverse shocks, often largely aggregate, when it is also costliest for individuals to share. Yet, how scarcity affects individual willingness to share and willingness to enforce sharing from others, an important ingredient in sustaining prosocial behavior, is little understood. This is what this paper examines. I conduct repeated within-subject lab-in-the-field experiments among Afghan subsistence farmers during a lean and a post-harvest season of relative plenty. These farmers experience seasonal scarcities annually. Using dictator and third party punishment games I separate individual sharing behavior from the enforcement of sharing norms. While sharing exhibits a high degree of temporal stability at both the aggregate, and, to a large extent, the individual level, the enforcement of sharing norms is substantially weaker during the

¹This work has been published in Bartoš, V. (2016) "Seasonal Scarcity and Sharing Norms", CERGE-EI WP series, No. 557. I thank Michal Bauer, Karna Basu, Jana Cahlíková, Alexander W. Cappelen, Lee Cronk, Dirk Engelmann, Ernst Fehr, Randall Filer, Peter Katusčák, Andreas Ortmann, Gérard Roland, Ondřej Rydval, Jakub Steiner, Bertil Tungodden and audiences at NHH Bergen, UC San Diego, NYU, Rutgers University, CERGE-EI, GDN (New Delhi), SEEDEC (Bergen), NIBS (Nottingham), IMEBESS (Toulouse), and ESA (Heidelberg) for invaluable comments. I am grateful for the hospitality and logistical support of the NGO People in Need, Afghanistan. I also thank Akram Rasaa, Mohibullah Mutahed, Kamran Shahzad and Yar Mohammad Rajabi for excellent research assistance. Financial support from GAUK (no. 46813), Czech Science Foundation (no. 13-20217S), and the GDN (RRC13+11) is gratefully acknowledged. All remaining errors and mistakes are mine.

lean season. The findings suggest that farmers are capable of sustaining mutual sharing through transitory periods of scarcity. It remains an open question whether exposure to unexpected shocks or prolonged periods of scarcity might result in the breakdown of prosociality due to loosened sharing norms enforcement on a community level.

1.1 Introduction

Sharing is a well documented source of informal insurance in village economies or poor communities where it frequently substitutes for formal insurance. It gains the uppermost importance during periods of scarcity; nonetheless, this is the period when sharing becomes most costly for those who share. This trade-off makes our understanding of how scarcity affects sharing an open question. To sustain sharing, societies require its members to punish shirkers (e.g., Boyd et al. 2003; Fehr and Fischbacher 2004a).² When the threat of punishment is missing, individuals disciplined to behave cooperatively start behaving selfishly, commencing the cycle of social erosion. Another question thus arises: Does scarcity affect individual willingness to engage in the enforcement of sharing? This paper addresses both these questions.

Previous research has examined *if* people share (Townsend 1994; Morduch 1995; Jalan and Ravallion 1999), *why* they share (List 2004; Leider et al. 2009; DellaVigna, List, and Malmendier 2012), and *who* they share with (Barr and Genicot 2008; Attanasio et al. 2012), or a combination of the three (Ligon and Schechter 2012), but so far there are only very few studies examining *when* to share and no economic studies that would causally determine a link between scarcity, and willingness to share and enforce sharing norms.

I establish this link by examining the sharing and norm enforcement behavior of

²Willingness to engage in costly third-party punishment in which materially uninterested individuals are willing to forego gains to punish unfair behavior has been documented in economic experiments (Fehr and Fischbacher 2004b; Bernhard, Fischbacher, and Fehr 2006) and was found to be positively correlated with the level of altruistic sharing in a cross-cultural study (Henrich et al. 2006). Fehr and Gächter (2000) show that cooperation can be sustained only if subjects have an opportunity to punish free-riders and gradually breaks down once the opportunity is removed, and, reassuringly, that cooperation can be restored once the enforcement mechanisms are reintroduced. The forms of punishment may range from physical attacks on non-cooperators, through gossip, all the way to ostracism of the non-cooperators (Maier-Rigaud, Martinsson, and Staffiero 2010). These forms of punishment are well documented in anthropology (Cronk, Chagnon, and Irons 2000), ethnography (Fessler and Navarrete 2004) and economic history (Greif 1993). Gülerk, Irlenbusch, and Rockenbach (2006) show that societies with punishment mechanisms are from an evolutionary perspective more competitive compared to societies where punishment mechanisms are lacking.

small-scale farmers in rural Afghanistan during a seasonal cycle of scarcity and relative abundance. The majority of the one billion people employed in agriculture in Asia and Sub-Saharan Africa are subsistence farmers dependent on highly volatile harvests, frequently affected by both aggregate and idiosyncratic shocks. This population is the most affected by seasonal scarcities (Sahn 1989; Devereux, Swan, and Vaitla 2008; FAO 2012; Khandker and Mahmud 2012).³ The cyclical nature of agricultural production, together with limited insurance, credit and savings markets, and low quality of storage technologies exposes many to seasonal scarcities (Basu and Wong 2015). Apart from seasonal migration (Bryan, Chowdhury, and Mobarak 2014), mutual willingness to share resources with others remains one of the few coping strategies.⁴ Since much of the world’s population is subject to agricultural cycles, it is of interest to learn how sharing concerns unfold at different points of the cycle.

A major challenge in examining sharing over time is that kinship, reputational concerns, reciprocity, or fear of retribution all confound the observed sharing behavior. Using observational data or narrative evidence, it is virtually impossible to distinguish between reputation-driven third-party punishment motivated by selfish motives from that driven by altruistic goals, not to say that quantifying social norms for cross-temporal comparison is inconceivable without using experimental methods. In order to overcome these issues, I conducted a controlled lab-in-the-field experiment using a one-shot dictator game (Idea originally used in Kahneman, Knetsch, and Thaler 1986) and a one-shot dictator game with a third party punishment option (Fehr and Fischbacher 2004b) examining the temporal stability of sharing behavior and of sharing norm enforcement among 207 subsistence farmers in northern Afghanistan.⁵ This remote rural society is exposed to dramatic aggregate and idiosyncratic seasonal shocks to consumption (NRVA 2008). I conducted two rounds of experiments with the same participants: one during the lean season and one during the post-harvest season. This provides me with a unique opportunity to inspect within-subject behavioral changes when exogenously exposed to more

³See Bryan, Chowdhury, and Mobarak (2014) for an extensive list of references documenting regular seasonal scarcities around the world.

⁴While food sharing is common in hunter-gatherer small-scale societies, the sharing of resources in more advanced communities may operate through the provision of informal loans on flexible interest rates with flexible repayment dates. Such behavior is frequently observed in poor communities (Collins et al. 2009).

⁵Economists and social scientists use dictator games to measure sharing motives (Camerer 2003). To address possible external validity concerns, Barr and Genicot (2008) show that risk-sharing decisions—namely risk-sharing network formation—observed in a similar experimental task reflect actual risk-sharing behavior in Zimbabwean villages.

or less scarcity.

Previous literature offers conflicting views as to whether sharing increases, remains constant, or decreases with resource scarcity. Moreover, to my knowledge, no study examining the effects of scarcity on other-regarding behavior differentiates between individual willingness to share and the willingness to engage in the enforcement of sharing norms. In other words, whether the behavioral change operates through the temporal instability of preferences or through a coordination problem on a community level.

Microeconomic theory would suggest that if the cost of sharing increases in the period of scarcity—which is plausible assuming concavity of the utility function over income or consumption—sharing behavior should decline.⁶ Yet experiencing scarcity also implies increased benefits to the receiver, who is more likely to be in need.⁷ Experimental, empirical, and theoretical literatures all give ambiguous predictions as to whether sharing or pro-social behavior in general increases or decreases during the period of scarcity.

On the one hand, Ostrom et al. (1999) argue that the scarcity of resources encourages more efficient institutional organization and enforcement mechanisms facilitating sustainable resource use. Anthropologists report narrative evidence of increased cohesion in both small- and large-scale societies facing seasonal food shortages (Evans-Pritchard 1969; Lévesque et al. 2000). Laboratory experiments support this by showing that extraction rates in a common pool game drop when resources become scarce (Osés-Eraso and Viladrich-Grau 2007).

On the other hand, scarcity is also shown to affect prosocial behavior negatively. Scarcity of common pool resources leads to more free-riding in ground water usage (Varghese et al. 2013) or in fisheries extraction rates (Maldonado, Moreno-Sánchez, and del Pilar 2009). Grossman and Mendoza (2003) show theoretically that common pool resources are extracted faster when survival is at stake. This is consistent with documented cases of increased selfishness during extreme food scarcities, such as famines or wars (Dirks 1980; Turnbull 1972). Scarcity further results in the general acceptance of loosened ethical behavior (Oster 2004; Miguel 2005), suggesting that social norms respond to temporal changes in the environment. Maritime disasters prove yet another setting in which social norms dissipate (Elinder and Erixson 2012). Less dramatic but equally important for the

⁶Andreoni and Miller (2002) show that a rising price of sharing indeed leads to a drop in sharing. Similarly, Fehr and Fischbacher (2003) conclude that with increasing cost of sharing, individuals become less willing to share in a dictator game or contribute to the public good in a public goods game.

⁷Engel (2011) shows in a comprehensive survey of dictator games that recipients' neediness increases amounts shared.

present study, Wutich (2009) shows that social networks' activity drops off during dry seasons. Changes in sharing norms under scarcity occur even in parent's treatment of their own children (Behrman 1988).

As for the importance of punishment behavior, groups ranging from small scale societies to large nation states are able to sustain cooperation if individuals are willing to engage in prosocial acts, together with enforcing prosociality from others, even against their own direct self-interest (Gintis 2000; Henrich and Boyd 2001; Boyd et al. 2003). Norm enforcement is especially critical in periods of shocks when the probability of group survival decreases, such as during wars, famines, disasters, or periods of scarcity, as in the case of this paper, when reputational motives are weak or non-existent. Enforcement reduces the proliferation of selfish types invading the population and thus increases prosociality.

Although the evidence of altruistic third-party enforcement of sharing in economic experiments is plentiful (Fehr and Fischbacher 2004b; Henrich et al. 2006; Bernhard, Fischbacher, and Fehr 2006), the literature examining its dynamics with environmental changes is scarce. Only Gneezy and Fessler (2012) get close by examining the dynamics of second-party enforcement of cooperation with the exposure to conflict. They show that enforcement intensified during the Israeli-Hezbollah war compared to a prior period or in the immediate aftermath. In their case the threat to the community came from an identifiable external threat. In the case of seasonal scarcity the threat comes from non-cooperative individuals within. Overall, the predictions as to whether scarcity is conducive or detrimental to sharing and its enforcement are unclear.

My findings are that despite substantial changes in income, consumption, health, and perceptions of stress within individuals across the lean and post-harvest seasons, sharing, measured by the amount passed in the dictator game, as well as in the third party punishment game, remain unchanged at the aggregate level and fairly stable at the individual level. However, the enforcement of sharing norms, measured by the willingness and the intensity of costly punishment of unfair allocations by monetarily uninterested third parties, are significantly weakened during the lean season. The drop in punishment of non-desirable behavior reflects a change in social norms rather than a shift in state-dependent individual preferences and can be attributed either to increased uncertainty about the intentions of others or to increased grievances suppressing the expression of altruistic punishment, with limited evidence favoring the latter. The observed results are quantitatively similar for two different groups represented in the study—one made up

of predominantly Sunni Tajiks and the other of predominantly Shia Hazaras—allowing for more generalizable statements about the findings presented. Although that I do not observe a change in dictators’ willingness to share across seasons it is plausible that during a prolonged period of weak enforcement under scarcity sharing behavior would drop. This is an established finding in laboratory experiments where prosocial behavior gradually deteriorates with unavailable enforcement mechanisms (e.g., Fehr and Fischbacher 2004a).⁸

There have only been a few experimental studies assessing the effect of scarcity on prosocial behavior. The present experiment is, to my knowledge, the first to examine the temporal stability of sharing in a setting where dramatic changes to consumption might possibly lead to changes in individual behavior. Second, it is the first paper examining the temporal dynamics of sharing norm enforcement using a third-party punishment game.

1.2 Related Literature

My paper speaks to different streams of literature:

First, recently a literature on the endogeneity of social preferences has been emerging. Social preferences have been found to be shaped in early childhood (Fehr, Bernhard, and Rockenbach 2008) through adolescence (Almås et al. 2010) and vary markedly across cultures (Henrich et al. 2010). All these studies examine long-term processes of preference formation, whereas the current paper analyzes possible dynamics over short-term periods of scarcity.

Second, conflict has been described as an important factor shaping human prosociality (Choi and Bowles 2007) and experimental studies confirmed the causal link between exposure to warfare and parochial altruism (Voors et al. 2012; Bauer et al. 2014). Parochialism induced by exposure to inter-group conflict differs from the scope of the present study in that war is an unexpected event in which the threat comes from outside of the society. The present study speaks to possible short-term effects of resource scarcity on sharing behavior. This also differs from recent studies examining effects of unexpected natural disasters on social preferences (Cameron and Shah 2015).

Third, the paper speaks to the emerging experimental literature examining temporal

⁸Similarly, Gneezy and Fessler (2012) link the increased willingness to punish in-group non-cooperators during wartime to the evolution of human cooperation, despite the fact that they do not observe any change in the ultimatum game transfers.

stability of preferences. Recent studies have shown that time preferences (Meier and Sprenger 2015), risk preferences (Andersen et al. 2008), and cooperative preferences (Volk, Thöni, and Ruigrok 2012) remain stable over time. However, all of the studies mentioned were carried out in stable environments of developed countries. My study is the first of its kind to provide evidence of temporal stability of sharing preferences in an environment exposed to substantial, yet to some extent expected environmental shocks.

Lastly, the paper speaks to the sparse literature examining temporal dynamics of social norms using economic experiments. To my knowledge, only Gneezy and Fessler (2012) examine changes in enforcement of cooperation during wartime.

The paper closest to mine is Prediger, Volla, and Herrmann (2014). They examine the effect of resource scarcity on cooperation and anti-social behavior among Namibian villagers using economic experiments in their natural environment where they are exposed to different levels of resource scarcity. The study shows that anti-social behavior is higher in the area exposed to higher scarcity of resources, but does not find any difference in levels of cooperation across the areas.⁹ Their study, however, differs from mine in several aspects. First, it does not differentiate between the role of individual prosociality and communal enforcement but rather concentrates on behavioral differences across communities in public goods and joy-of-destruction games. Second, their study considers differences in behavior across two locations exposed to different environmental conditions in a long term, while my study examines short-term effects of scarcity on cooperation within a particular community, with villagers participating repeatedly in an experiment when their environmental conditions are changing exogenously.

Another closely related study is that of Fisman, Jakiela, and Kariv (2015). They examine the effect of the 2008 Great Recession on sharing behavior to find that people become more selfish after experiencing an economic downturn, both actual and a lab-induced. Their study differs from mine in the following: First, it considers the sharing and not the enforcement part. Second, the study is conducted in a developed country where recession might trigger different responses than in a developing country. Third, their study examines behavior of different groups of individuals, rather than observing the same individuals over time as I do in this paper.

The method I employ resembles that of Mani et al. (2013) who examine the effect of scarcity on cognitive abilities in a population of Indian sugarcane farmers. Mani et al. (2013) observe their participants over the pre- and post-harvest seasons and compare

⁹The cooperation behavior results are only reported in an earlier working paper.

the results before and after. Similarly, Behrman (1988) studies temporal dynamics in parental preferences over nutritional allocation between sons and daughters with exposure to seasonal scarcities following the same sample of rural Indians. The present study aims to contribute to this stream of literature by examining the temporal stability of sharing and of sharing norms enforcement in a highly volatile environment of Afghanistan.

1.3 Experimental Design

1.3.1 Sample Selection

The participants were recruited for the experiments in 10 randomly selected villages in the Marghzar and Amrakh areas of Zari district in Balkh province, northern Afghanistan, a remote area at high elevation. With more than 60 percent of the population living below the poverty line, Balkh is one of the poorest provinces in Afghanistan (NRVA 2008). The vast majority of the local population subsists on agricultural production or agricultural labor. We invited all land-owning farmers, a maximum of one adult person per household was allowed. The head of the household—the main bread winner—was strongly preferred. Due to cultural constraints, only males were invited.

To answer the question whether sharing and enforcement of sharing norms vary with exposure to resource scarcity I exploit the fact that farmers in this area face annual seasonal food shortages. I conducted 20 experimental sessions in 10 villages with 291 adult male farmers in the lean season of April 2013 and an additional 20 sessions in the same villages with 207 participants who we managed to contact also in the post-harvest season in October 2013. To overcome possible “calendar effects”, I conducted the experiments outside of major Islamic holidays, harvest time, or bazaar days and no significant events were reported when we conducted the experiments.

In the post-harvest season we also recruited an additional 212 new participants to substitute for the 84 participants who dropped out and to provide a sample of “virgin” participants who participated only in the second round to control for potential order effects. The selection procedure was the same as in the lean season round, and the differences between the samples presented in Table 1.A.2 are discussed in subsection 1.5.4. Despite some differences between the respective samples in observable characteristics, I show that the behavior in games does not differ across samples and importantly, in the main analysis I only focus on the behavior of those, who participated in both rounds. Each

session was conducted with 12 or 15 participants. The participation in each experimental round was voluntary and the participants could leave at any time. All participants decided to complete the tasks within each round.

Demographic characteristics for the sample of the 207 participants participating in both experimental rounds are presented in Table 1.1. Half of the sample are Sunni Muslims (51 percent) mainly of Tajik ethnic origin and the other half is Shia Muslims of predominantly Hazara ethnic origin, all living in almost perfectly segregated areas.¹⁰

It is important to note that 84 subjects who participated in the first experimental round did not participate in the second experimental round. Out of them 62 (74 percent) migrated either to Iran, to Mazar-e-Sharif, Kabul, or to another village for seasonal work. Only the remaining 22 (26 percent) did not show up either because of working elsewhere at the time of the experiment, being sick, or attending a wedding at the time of the assigned experimental session. Reassuringly, no one declined to participate due to reasons related to the experiment.

Note that the selective attrition would systematically bias the results only if it were correlated with the stability of sharing and with willingness to engage in third-party norms enforcement.

1.3.2 Seasonal Effects

There is vast evidence that farmers in developing countries are exposed to substantial fluctuations in incomes and consumption over the year (Devereux, Swan, and Vaitla 2008; Khandker and Mahmud 2012). Table 1.2 presents the seasonal differences in observable characteristics among the sample of participants in both seasons. The data show that seasonality does indeed matter. The participants' average monetary income per OECD equivalence scaled household member in the previous month in the lean season is only 69 percent of the post-harvest season income. Also, 59 percent of participants reported having no monetary income in the lean season compared to 38 percent of participants in the post-harvest season. Smoothing consumptions with own income across seasons is unlikely due to almost non-existent savings in the area.

Meat is consumed less frequently during the lean season. The share of people in debt increases from 70 percent in the post-harvest season, already high, to 86 percent

¹⁰I do not control for religion in the analysis because individual religious affiliation is perfectly correlated with village affiliation. I use village fixed effects in regressions that thus control for possible effects of religion too.

in the lean season. The participants also seem to have less money available for lending out during the lean season as the share of subjects lending money to others decreases from 39 percent in the post-harvest season to 29 percent in the lean season.¹¹ Further aggravating the severity of the lean season, the participants report being much more likely to be unable to work due to injury or illness, they feel generally more stressed, and are affected by shocks such as crop pests and diseases, livestock diseases, as well as human diseases. Irrespective of the season, 25 percent of the participants report that someone from their household has been out of the village, migrating for work.

Figure 1.1 shows that the participants are well aware of the seasonal swings over the year. Responding to a question to select three months of a year that are generally most difficult for them to cope with and three months of a year that are generally least difficult for them to cope with, most participants perceive the winter and the spring months (the lean season) as the most difficult to live through and the summer and the autumn months (the harvest and the post-harvest season) as the best months of a year.

1.3.3 Experimental Tasks

Each experimental session consisted of two tasks. A one-shot dictator game (DG; Idea originally used in Kahneman, Knetsch, and Thaler 1986) and a one-shot dictator game with a third party punishment option, the third-party punishment game (TPPG; Fehr and Fischbacher 2004b; Bernhard, Fischbacher, and Fehr 2006). To control for order effects I randomly manipulated the order of tasks. The participants were rematched after each task and across rounds in order to avoid strategic behavior and possible reciprocal concerns. After the experiment each participant was surveyed.

The DG allows me to examine the temporal stability of individual sharing behavior in the absence of confounds of kinship, reciprocity, reputation building or the fear of social sanctioning for non-desirable behavior. In this quasi-game a dictator, Person A (PA), divides a given endowment (10 experimental currency units, ECUs) between himself and a passive receiver, Person B (PB). PB is also one of the participants in the same experimental session as PA, but he receives no endowment and only learns the final allocation of money. The game allows for 11 strategies, as only whole units can be passed. The allocation depends entirely on PA's own willingness for unconditional sharing under the veil of anonymity, as his identity is never revealed to PB. Thus, the individual is

¹¹As other studies from developing countries have found, many people are lenders and borrowers at the same time (Collins et al. 2009).

motivated to reveal his true sharing preferences. For simplicity, the ECUs in the game are represented by money slips evoking 20 AFN banknotes, not by real money. The conversion rate is 1 ECU = 20 AFN.

In order to test the temporal stability of sharing norm enforcement, I administer a TPPG. The game allows a monetarily uninterested third party—Person C (PC)—to observe the sharing behavior of a dictator—PA—in a DG where even PA and PB are aware of PC’s presence. First, PA decides how much of the 10 ECUs of his endowment to pass to PB who has no endowment as in the DG described earlier. PB only learns PA’s final decision and has no control over it. Second, PC may decide to punish the dictator for his behavior but only at a cost to himself. Each PC is endowed with 5 ECUs and he can either refrain from punishment or pay 1 or 2 ECUs to subtract 3 or 6 ECUs of PAs payoff, respectively. This distribution ensures that in a situation when PA behaves as an egalitarian and PC decides not to punish such behavior, all players leave the experiment with 5 ECUs. However, PCs do not observe PAs’ actual behavior. Rather, I elicit their reaction to all possible behaviors of PA using a strategy method.¹² PC’s willingness to pay to punish provides me with a direct measure of willingness to engage in altruistic enforcement of specific sharing norms. The variable of interest is the minimum acceptable PA offer to PB that is not punished by PC. Further in the text, I denote the minimum acceptable offer as MAO (originally used in Henrich et al. 2006).

1.3.4 Experimental Procedures

The experiments were announced one day in advance. The villagers were informed that an experiment requiring a commitment of four hours of their time will be conducted in their village for which they will earn at least 100 AFN (approximately 2 USD) as a show-up fee, but possibly more.¹³ All interested farmers were gathered in a community center (a

¹²Brandts and Charness (2011) survey 29 studies that directly compare the strategy method to direct-response elicitation. While in the majority of cases no difference between the two methods is found, the only exception is games with punishment. Out of four studies including a punishment option surveyed, three observed lower levels of punishment when the strategy method was used. Reassuringly, in all cases the treatment effects were detected using both methods and the effects were in the same direction.

¹³An average daily wage is 150 AFN, but it is not possible to find work every day in the area. During the off-season work is particularly scarce. Importantly for my study, the size of the initial endowment does not seem to influence the relative transfers in dictator games (Engel 2011) or punishment games (Kocher, Martinsson, and Visser 2008) to the extent that might invalidate the results of the present study. In order to validate this claim, I conducted several experimental sessions with stakes increased by 50 percent in the 2013 lean season only to find that the relative transfers do not differ from the transfers in games with the original endowment size. The 50 percent increase reflected the reported 50 percent increase in prices of most common consumption goods during the lean season compared to the

guesthouse, mosque, or a village leader’s house) the morning just before the first session. If more villagers showed up for an experimental session than we could accommodate, we either invited them for another session if there was one conducted in the same village or we ran a lottery in which we selected the participants by chance. Consequently, the actual participants randomly picked an ID number, which determined their role in the experiment. The numbers of participants by role, village, and round are reported in the Appendix, Table 1.A.1.

As is common in economic experiments carried out with low-literacy subjects, the instructions were first explained in a group using practical examples and visual aids (See Figure 1.A.1), and only then were the actual experiments carried out with the subjects individually (See Figure 1.A.2).¹⁴ Before making their actual decisions, all participants were shown several examples, were allowed to practice several scenarios themselves, and were then asked to answer several control questions. The research assistants explained the task until the participants fully understood and the experiments were carried out only after participants’ full comprehension. Only one participant failed to pass the comprehension test due to hearing problems, not the inability to comprehend the task. The instructions were presented orally in the local language, Dari, and were translated back to English.

Communication in all rounds of experiments was not allowed and all tasks were strictly anonymous. Only one task was randomly selected for the payment to avoid strategic play across experiments. This procedure was revealed to the participants in the instructions.

Although the participants received their payments at the end of each experimental session they did not receive any feedback on their actions and the actions of other players. Average earnings were about 190 AFN including the show-up fee (100 AFN), which is slightly above the average daily wage of a casual laborer. In order to prevent post-play retaliation, all payments were carried out in private and this was communicated to the subjects before the play.

post-harvest season. Results are available upon request.

¹⁴The instructions and procedures I used are inspired by Bernhard, Fischbacher, and Fehr (2006) and by Henrich et al. (2006). Instructions are available in the Appendix 1.A.2.

1.4 Results

In this section I first discuss both aggregate and individual-level temporal stability of sharing behavior. Then I present the behavioral change in willingness to enforce sharing norms over time. In the main results discussed in this section I restrict the sample to farmers who participated in both rounds.

1.4.1 Temporal Stability of Sharing Behavior

I begin by discussing the temporal stability of sharing behavior. First, I present the aggregate results of sharing behavior. Second, since the design of the experiment allows me to observe the sharing behavior within the same individual across seasons, I present the results on the within-subject stability of sharing.

Does the aggregate sharing behavior differ across seasons? Columns 1 and 3 in Table 1.3 show that in the DG the PAs transferred on average 3.03 ECUs to PBs in the lean season compared to 3.22 ECUs in the post-harvest season, the difference being statistically insignificant (Mann-Whitney U-test, MWT: $p=0.48$, $n=136$). Similarly for the TPPG, I find that an average transfer of 2.87 ECUs in the lean season and 3.10 ECUs in the post-harvest season, the difference being again statistically insignificant (MWT: $p=0.41$, $n=136$).

I test the temporal stability of sharing behavior using the following regression model:

$$T_{it} = \alpha + \beta LS_i + \gamma X_{it} + \varepsilon_{it} \quad (1.1)$$

where T_{it} is the amount passed by the individual i in the experimental game in the period t , which is either the lean season or the post-harvest season. The α is an intercept, LS_i is the treatment variable equal to 1 in the lean season, X_{it} is a set of individual characteristics¹⁵, and ε_{it} is the error term.

¹⁵In the main estimations I either omit the control variables, add only a set of village dummy variables (with the dummy for village Kalakhan-e-Bala is excluded to avoid perfect multicollinearity), or add both village dummy variables and individual level characteristics such as age, number of years in school, number of individuals living in the individual's household, individual's income in the previous month, and the comprehensive poverty index proxy. The poverty index at a given point of time is estimated using the principal component analysis. The 1st principal component of each poverty measure for a given season is constructed using current individual income, animals owned, assets owned, variability of food consumed, meat eaten in a given week, days unable to work due to illness or injury in the previous month, a short version of the perceived stress score (Cohen, Kamarck, and Mermelstein 1983), and dummy variables representing unusual health shocks to humans, animals, and plants. Note that the results presented in this paper are robust to the use of different sets of controls (additional analysis available upon request).

Table 1.4 shows that the behavior across seasons remains stable both in the DG and the TPPG when using a regression framework. The first model (Columns 1 and 4 in Table 1.4) does not include any controls. The second model (Columns 2 and 5) controls for village-specific effects, as the village fixed effects explain about 16 percent or 13 percent of the variance in the DG or the TPPG transfers, respectively (See Table 1.A.3). Finally, the third model (Columns 3 and 6) further controls for additional individual level controls. In neither case is the variable *lean season* statistically significantly different from zero and we can conclude that the sharing behavior does not change across seasons for either the DG or the TPPG.¹⁶

Figure 1.2 examines the cumulative distributions of respective amounts transferred in the DG (Panel A) and the TPPG (Panel B) across the two seasons. Apart from the difference in the frequency of PAs sending 3 ECUs both in the DG (difference in frequencies across rounds borderline significantly different from zero, $p=0.09$) and the TPPG (marginally insignificant, $p=0.13$), the distributions are identical, a necessary condition for stability of preferences. The Epps-Singleton Two-Sample Empirical Characteristic Function (ESCF) test cannot reject the equality of distributions for neither the DG ($p=0.22$), nor the TPPG ($p=0.34$).¹⁷

Finding 1: *On the aggregate level I find that the sharing behavior in the DG and the TPPG does not vary with short term exposure to scarcity.*

I now turn to the analysis of the within-subject stability of sharing. In total, we successfully tracked 68 PAs. These participants were exposed to the same experimental procedure in both the lean season and in the post-harvest season, six months later.

I examine the correlations in sharing behavior across seasons and individual changes in sharing behavior. I compare the actual changes in sharing behavior to a reference situation in which I treat the distribution of transfer choices as randomly allocated across

¹⁶In the main regressions I use the commonly reported OLS with clustered standard errors. The results are robust to using ordered probit, which takes into account the discrete nature of the dependent variables. See Tables 1.A.4 and 1.A.5 for the replication of OLS results.

¹⁷The distribution of DG transfers fits between the classifications of the developing country and an indigenous society subject pool classification used in the DG meta study by Engel (2011). The Afghan PAs are much more likely to pass positive amounts to PBs than the Western subjects (91 percent versus 67 percent in the Western societies, 81 percent in the developing countries and 95 percent in the primitive societies), slightly less likely to pass equal share (21 percent versus 20 percent Western, 27 percent developing and 28 percent primitive societies), but no one in this sample passes the entire pie unlike 5 percent of the Western subjects and 1 percent both in developing countries and in primitive societies. Similar comparison for TPPG transfers is not possible, since the game has not been used so extensively and no effort to conduct a meta-analysis has been made.

individuals. First, I describe the stability of sharing behavior in the DG and then I comment on the stability of behavior in the TPPG.

Panel A of Figure 1.3 presents the histogram of *changes* in individual behavior in the DG, specified as a difference between the lean and the post-harvest season transfers. It reveals that more than 30 percent of individual decisions in the DG remained constant across both seasons. Moreover, almost 65 percent of decisions remained within a change of one ECU or 10 percent of the PAs endowment. The correlation between DG transfers in the lean season and in the post-harvest season is 0.52 ($p < 0.01$). Such stability is relatively high compared to other studies examining temporal stability of preferences.¹⁸

It is possible that the result presented here as a proof of temporally stable sharing behavior could arise as a confound, and would arise even if the DG choices were drawn randomly. We can rule out this possibility, as each choice from the entire set of possible transfers would have to be represented uniformly, which is clearly not the case without any need for statistical testing. On the other hand it is well plausible that due to the limited choice space observed in the cumulative distribution of choices in Figure 1.2 with the majority (75 percent) of PAs transferring between 2 and 5 units, it could be that the temporal stability of the sharing behavior is an artefact of the experiment. In order to rule out this possibility, I conduct an exercise in which I randomly assign choices from the set of all realized transfers in the post-harvest season to PAs. After reshuffling the PA choices 10000 times, the average number of equal choices across both seasons is around 15.6 percent, and 42.5 percent of decisions remain within a change of one unit, much lower than the actually observed values.

Next, I discuss the within-subject stability of TPPG results. Although statistically significant ($p = 0.07$), the correlation of individual behavior in the TPPG across seasons is 0.22, much lower than the correlation discussed in case of the DG. Yet even such correlation would be generally accepted as fairly stable over time in the psychological literature (see footnote 18). Panel B in Figure 1.3 shows that only 13 percent of individuals sent equal amounts in both seasons, even though the share of individuals with changes within a margin of one unit reaches over 55 percent.

In a similar exercise as presented for the DG, I simulate what would have happened had

¹⁸Literature in psychology examines the stability of preferences in much more detail than economics does. Surveys examining stability of single cross-situational measures usually report temporal stability in a range between 0.2 to 0.3 (see e.g. Block 1983; Jessor 1983) and perceives such correlations as indicating relatively stable preferences, while within this interval. Similarly to my findings, Meier and Sprenger (2015) report a correlation of 0.5 in individual time preference choices in an experiment repeated twice over a year with the same set of subjects and label such correlation as high.

the distribution of TPPG transfer choices been randomly drawn from the distribution of choices in the post-harvest season to see how many individuals would have sent equal split in such hypothetical case. The average share of participants sending equal amounts in both seasons after random reshuffling in 10000 repetitions is over 16 percent. This implies that the results I obtain in my experimental data are no better than due to random chance. More reassuringly, conducting the same exercise for the variable indicating a transfer difference within a margin of one ECU, the share is about 43 percent, indicating some degree of individual stability within this extended margin.

***Finding 2:** Transfers in the DG are temporally stable within individuals, suggesting stability of sharing. To a lesser extent I also observe within individual temporal stability in TPPG.*

1.4.2 Temporal Stability of Norm Enforcement

Now I analyse the behavior of PCs in the TPPG in order to understand the dynamics of sharing norm enforcement with exposure to scarcity of resources. I first discuss the aggregate punishment results using the sample of farmers who participated in both rounds, and I examine the within-subject results later.

Figure 1.4 shows the distributions of PCs' minimum acceptable offers in the TPPG (MAO) in both the lean and the post-harvest seasons. MAO is the lowest PA's transfer to PB that a PC would accept.¹⁹ For example, if a PC decided to engage in either type of punishment of the PA for sending anything less than or equal to 2 ECUs to PB, then the MAO for this PC is equal to 3 ECUs. The lowest value for MAO is 0 ECU if PC decides not to punish any kind of PA's behavior. I was able to elicit MAO for 60 out of 71 PCs in the lean season (85 percent) and for 63 out of 71 PCs in the post-harvest season (87 percent).²⁰ The subjects for whom I am unable to construct MAO behaved in an inconsistent way, punishing transfers largely at random without any systematic pattern. In the analysis below I use the 123 valid observations.

Figure 1.4 shows that the Afghan participants in the role of PCs were willing to engage in costly punishment of PAs who were not willing to share enough. Regardless of season,

¹⁹In this text I do not differentiate between the intensity of punishment, but the results presented would only be strengthened by accounting for it. These results are available upon request.

²⁰In terms of the task comprehension, this makes my sample comparable to that of Henrich et al. (2006), who were able to assign MAO to 92 percent of their sample.

the probability of punishing PAs increases with PAs' transfers approaching zero.²¹

Unlike in the case of PAs' transfers, the punishing behavior of PCs is not temporally stable. Figure 1.4 shows that there is a significant decrease in the willingness to punish low offers from the post-harvest to the lean season. Speaking about magnitudes, PCs in the post-harvest season were on average not punishing offers equal to 3.03 ECUs and higher, while in the lean season the average MAO dropped significantly to 1.35 ECUs (Columns 1 and 3 in Table 1.3), reaching the levels of average transfers in the DG and TPPG. The difference in MAO across rounds is highly statistically significant (MWT: $p < 0.01$, $n=123$). I can also reject the equality of MAO distributions over time (Epps-Singleton, $p < 0.01$).

Table 1.5 shows that the increase in willingness to punish remains highly significant and of a similar magnitude even in a regression framework. Again, I use the model specified in Equation 1.1 where the T_{it} now stands for the MAO by individual i in time t . In the first model I do not control for any additional characteristics (Column 1 in Table 1.5), in the second model I control for the village level fixed effects (Column 2), and in the third model I control for both the village level fixed effects and the individual level characteristics together (Column 3). In all specifications MAO remains statistically significantly lower in the lean season round.

Importantly, the behavior of PCs is also reflected in beliefs of others. Apart from the main experimental task, I also measured beliefs using several incentivised questions. Regarding the punishment, I asked the participants whether they believe that most PCs in the current experimental session would punish a PA who decides to transfer zero ECUs. The results are presented in Table 1.3. Although insignificantly, the beliefs of PBs (lean season 68 percent vs. post-harvest season 78 percent; MWT: $p=0.18$, $n=136$) match the actual behavior of PCs and is of similar magnitude as beliefs of PCs about other PCs' willingness to punish zero transfers in their experimental session (lean season 65 percent vs. post-harvest season 79 percent; MWT: $p=0.06$, $n=142$). This suggests that the behavioral change across seasons is more generally considered in the population and is not just an artefact of the experiment among the group of PCs. This conclusion has to be taken with some caution, since the PAs beliefs do not match that of PBs and PCs. Unlike the other participants, the PAs expect the punishment of zero transfers in 70

²¹Such a pattern emerges even if we include the inconsistent punishers (analysis available upon request). Similarly, the results remain qualitatively unchanged in a more restrictive analysis that excludes an individual from the sample if he behaved inconsistently in either period.

percent of cases regardless of season (MWT: $p=0.89$, $n=135$).

As in previous studies (Fehr and Fischbacher 2004b; Bernhard, Fischbacher, and Fehr 2006; Henrich et al. 2006), the Afghan farmers are willing to engage in costly altruistic punishment for which they have to give up 20 percent or 40 percent of their endowment to punish non-desirable behavior. In terms of daily incomes, the amounts are equal to giving up 13 to 26 percent of average daily incomes to discipline others, a substantial amount given the tight budgets of the population studied. Overall, 93 percent of the PCs for whom I am able to construct the MAO are willing to punish a PA who decides to keep everything in the post-harvest season, a number comparable to the most punishing societies in the study of Henrich et al. (2006), the Kenyan Gusii and Maragoli tribes. This share drops to 62 percent in the post-harvest season, similar to the average punishment choice frequency for zero transfers in the 15 small-scale societies studied in Henrich et al. (2006) (MWT: $p<0.01$, $n=123$; Columns 1 and 3 in Table 1.3).

***Finding 3:** Afghan farmers substantially decrease intensity of norm enforcement mechanisms during the lean season.*

As in the case of the sharing behavior, the experimental design also allows me to examine punishing behavior across seasons within an individual. There were 52 PCs for whom I could construct the MAO in both rounds. The remaining 19 PCs behaved inconsistently in either of the seasons, but never in both. In the lean season 11 PCs behaved inconsistently compared to 8 PCs in the post-harvest season. Overall, 34 PCs decreased the level of punishment in terms of MAO between the post-harvest and the lean seasons, 5 PCs punished exactly the same across both seasons, and 13 increased the level of punishment. Figure 1.5 presents a histogram of individual changes in MAO across seasons.

As discussed, punishment decisions were elicited using the strategy method. Footnote 12 shows that earlier studies have found differences in punishment behavior when comparing the strategy method and direct-response elicitation. Taking these studies as a guide, we might assume that actual punishment decisions could be higher than those presented here. The reason for this is impulsive behavior when *hot*, directly elicited decisions are made. Further, reassuringly, the treatment differences in earlier studies have been directionally the same, regardless of the method used. There is a concern that the differences between the strategy and direct elicitation procedures might be different

across seasons. Mani et al. (2013) show that cognitive abilities decrease with exposure to scarcity during a lean season. It is plausible that lower cognitive ability is linked to more impulsive behavior (Kahneman 2011). Hence, we might expect to observe more impulsive behavior leading to increased punishment in the lean season even in the strategy method decisions. But it is the opposite that we observe.

What characteristics explain the behavioral change? Table 1.6 shows that regressing the difference in MAO between the post-harvest and the lean season on a set of regressors that include participant's age, years of schooling, number of household members, individual income in either of the seasons, or the poverty index in either season does not provide us with any explanation for the observed change in behavior.

1.5 Discussion

In this section I provide some evidence that the drop in the willingness to punish in the lean season can either be attributed to higher uncertainty about the intentions of others or due to higher wealth inequality present at that period. Also, I show that the drop in punishment is not determined by individual severity of the seasonal shock, but rather by the severity of the aggregate shock on the community-level. This implies that a communal social norm is driving the behavioral change. I also show some evidence that speaks for the generalizability of the observed behavior. Besides that, I rule out several possible caveats such as the role of order effects or the effect of changing marginal utility of wealth across seasons as possible explanations for the behavior observed.

1.5.1 Determinants of Seasonal Changes in Norm Enforcement

In Section 1.4.2 I show that there is a substantial drop in punishment behavior in the lean season compared to the post-harvest season. What factor is driving the difference? Several possible explanations can be put forth:

First, punishment might be perceived as a normal good, demand for which increases with increasing income. Examining the correlation between MAO and individual income (Column 3, Table 1.5), I actually find an opposite: a negative correlation ($\beta=-0.61$, $p=0.03$). This effect may be driven by the fact that the wealthier individuals are in general less likely to engage in altruistic punishment. The nature of the data also allows me to examine the *change* in income within an individual across seasons. Comparing

the MAO for those PCs whose reported income was higher in the post-harvest season compared to the lean season ($n=21$) and those whose income did not increase in the post-harvest season ($n=31$), I find that MAO is not significantly statistically different across these groups (MWT: $p=0.42$, $n=52$).²² Specifically, the *change* in MAO for those whose income did not increase between the post-harvest and the lean season is equal to -1.74 , while the change in MAO for those whose income increased is -1.14 . Monetary income might, however, not be the best proxy of wealth in agrarian communities. Conducting a similar analysis for the seasonal difference in the comprehensive poverty index yields similar results. Importantly, the number of PCs whose poverty index was lower in the lean season compared to the post-harvest season is 10, while for the remaining 42 PCs the poverty index increased. Income effects thus do not plausibly explain the observed drop in sharing norms enforcement in the lean season.

Second, Grechenig, Nicklisch, and Thöni (2010), Xiao and Kunreuther (2015), and Bornstein and Weisel (2010) find that the punishment level drops with rising uncertainty about PA's intentions. It is plausible that increasing uncertainty about the PA's financial situation might cause the lower punishment levels observed in the lean season. In other words, the PC in the lean season cannot differentiate between a selfish and a needy PA, which is the reason why he rather abstains from getting involved in the judgment and possible later regret if he decided to punish a needy individual. This uncertainty is generally higher in the lean season. Not only is income level generally lower, leaving more people below the subsistence threshold,²³ it is also much more variable. The Gini coefficient for the entire sample reaches 0.47 in the lean season and drops down to 0.33 in the post-harvest season. Table 1.2 (Columns 2 and 4) shows that the standard deviation for individual income is significantly higher in the lean season (Variance ratio test: $p<0.01$, $n=278$). Similarly, the standard deviation of the comprehensive poverty index is also significantly higher in the lean period (VR test: $p<0.01$, $n=278$). However, the predictive power of a model regressing the seasonal change (both individual and village-average) in willingness to punish on the average village-level *variance* of the poverty index or of

²²The number of observations in this analysis is 52. This is the number of subjects for whom I was able to construct the MAO in both rounds. The income of 14 PCs remained constant across seasons and for 17 PCs it increased in the lean season compared to the post-harvest season. However, while median income was 2500 AFN higher in the lean season for the group of PCs whose income increased, the median drop in income for the group of PCs whose income decreased in the lean season was 4500 AFN.

²³NRVA (2008) reports that the food consumption of 48 percent of rural Afghans is below a poverty line during the lean season, compared to 21 percent in the post-harvest season.

income is very small.²⁴ Nevertheless, the small sample size of only ten villages does not allow us to rule out the proposed hypothesis.

Further, it can be argued that the PCs might expect the PAs to overcome uncertainty about the neediness of PBs by keeping the money from the experiment and sharing it afterwards in person. However, none of the participants reported willingness to share the money with anyone outside of his family in a post-experiment survey. Nearly 90 percent of participants in the lean season and over 96 percent of the participants in the post-harvest season reported that they plan to spend the money from the experiments on food or other household expenses.

Third, increased inequality during periods of scarcity has also been shown to predict the rise of grievances, which is one explanation for the rise in conflicts during scarcity (Hidalgo et al. 2010; Hsiang, Burke, and Miguel 2013). It is possible that increased acceptance of violence in solving problems can be associated with the observed decrease in willingness to punish non-cooperative behavior during the period of scarcity. In my sample I observe an increased number of individuals who were engaged in disputes²⁵ during the lean season when compared to the post-harvest season (14.5 percent versus 7.7 percent; MWT, $p=0.02$, $n=414$).

Table 1.A.6 presents supportive evidence for the role of increased grievances in explaining the drop of punishment. The regressions show a negative correlation between the change in the average village-level share of individuals engaged in a dispute between the lean season and the post-harvest season and the change in MAO between the lean season and the post-harvest season. The first three models use average village level change in MAO as a dependent variable. Despite the small number of observations—ten villages—the effect is highly significant in all three regression specifications that use different analytic weights. Although significance is lower, models 4 to 6 show effects of similar magnitude using individual level changes in MAO as a dependent variable. A simple back of the envelope calculation, with the average change in the share of individuals engaged in disputes being less than 7 percentage points, suggests that the estimate explains around 30 percent of the observed change in punishing behavior.²⁶ It is important to note that this effect cannot be interpreted causally. Despite that, the link between relatively higher

²⁴Results are available upon request.

²⁵Individuals were asked a question whether they were "engaged in a dispute in the previous four weeks".

²⁶Conducting a similar analysis on an individual level rather than on a village level does not yield significant estimates.

engagement in disputes and relatively lower punishment behavior in the lean season on a village level is telling.

The design of the experiment does not allow to separate the second and third explanations. One way or the other, Fehr and Fischbacher (2004a) have provided strong evidence replicated in numerous experiments that without norm enforcement mechanisms groups gradually dwindle to a non-cooperative equilibrium. Boyd et al. (2003) provide a theoretical model showing that third party punishment helps societies to maintain cooperative equilibria even in larger groups and its absence leads to a collapse of cooperation, as selfish individuals invade the population and their behavior provides them with higher payoffs compared to the payoffs of cooperators. A cross-cultural study shows evidence of positive correlation between altruistic sharing and sharing norm enforcement (Henrich et al. 2006). Thus, regardless of PCs' motivations, the drop in norms enforcement in the lean season increases the likelihood of a drop in sharing.

On the other hand, I do not observe a change in behavior of PAs in the TPPG, which speaks against the claim that sharing deteriorates with the lack of norm enforcement. But prosocial behavior both in Boyd et al. (2003)'s theoretical model as well as in Fehr and Fischbacher (2004a)'s experimental study deteriorates only gradually, as the selfish types start invading the population. My result is consistent with such gradual deterioration of cooperative behavior in the case of prolonged scarcity of resources of which—by playing a one-shot game—I only observe the initial stage and of which Hsiang, Burke, and Miguel (2013) (emergence of conflict due to climatic change) or Dirks (1980) (breakdown of cooperation during famines) observe the final stage. Similarly, Gneezy and Fessler (2012) do not observe a change in behavior of PAs in the ultimatum game from peacetime to wartime played only once in each period, despite the observed increase in punishment behavior during wartime.

Similarly to Wutich (2009) who documents that weakening of social networks is only temporary for the duration of a dry season and returns to original levels with the end of the dry season, Afghan farmers maintain some stabilizing mechanisms that prevent them from plunging into non-prosocial equilibria. However, it seems that they lack mechanisms preventing the collapse of cooperation in times of prolonged scarcity or of unexpected shocks. This might explain the dynamics of collapse of cooperation during famines (Turnbull 1972; Dirks 1980; Ravallion 1997). As my results suggest, the drop in prosocial behavior observed in this literature does not necessarily stem from changes in individual preferences, but rather from weaker social norm enforcement. The aim of the next section

is to that the change in altruistic punishment can indeed be attributed to changing social norms, rather than to mere individual preferences responding to changing individual conditions.

1.5.2 Individual Preferences or Social Norms as Determinants of Punishment Behavior

Although resorting to punishment in the TPPG is generally understood as an expression of willingness to sustain social norms (Henrich and Boyd 2001; Boyd et al. 2003; Fehr and Fischbacher 2003; Henrich et al. 2006), the behavior could also be driven by state-dependent individual other-regarding preferences. Such behavior would be consistent with models of inequality aversion (Fehr and Schmidt 1999) or the theory of reciprocity (Falk and Fischbacher 2006), in which an unkind act of a PA towards a PB has a negative effect on PC's utility. Note that individuals' wealth and his expectations of the wealth of others also have to be incorporated into the model's parameters. The act of punishment in such models would have two effects: first by the effect of deterrence preventing PAs to engage in unkind behavior in the first place and second by the moderation of selfish PAs' advantageous inequality by reducing their payoff relative to that of other players.

Understanding this distinction is important. If social norms guided the observed behavior, moral authorities in the society could have their say in affecting individual behavior. On the other hand, individual level interventions would hardly make any change, at least on a short term horizon. The opposite argument can be made if individual preferences were driving the observed behavior. My data speak in favor of the explanation based on social norms:

The regression models including variables representing individual exposure to scarcity—income and the comprehensive poverty index in either season—cannot fully explain behavioral change in punishment behavior (Table 1.6). Splitting the results by whether a particular PC perceived himself as relatively richer or poorer compared to his fellow villagers in the given season also does not predict the behavioral change (results available upon request). On the other hand, examining the average seasonal change in exposure to scarcity within a village is linked to the change in the TPPG MAO between the lean and the post-harvest season in a way that would support the norms-based explanation: the more severe the shock in the average village-level poverty, the larger the drop in MAO. Table 1.A.7 summarizes the results both using the average village-level change in MAO

as a dependent variable (models 1 to 3) as well as the individual-level change in MAO as a dependent variable (models 4 to 6).²⁷ It thus seems more plausible that the observed drop in punishment during the lean season is driven by changes in social norms.

1.5.3 Generalizability

Even though more research needs to be done in understanding whether the presented results can be generalized to other populations, it is important to point out that the results are valid for two very different groups. As shown in Table 1.1, half of the sample in my experiment are ethnic Tajiks and the other half are ethnic Hazaras, the second and third largest ethnic groups in Afghanistan respectively. While the former are Sunni muslims, the latter are Shia muslims, a minority in Afghanistan.

Tajiks are of Persian origin. They are, after Pashtuns, the second largest ethnic group in Afghanistan with around 32 percent of the population. In the Balkh province where the experiments have been conducted Tajiks are the predominant ethnic group, with around 44 percent of the population (DHS, 2010). The governor of the province is a Tajik himself. Hazaras, people probably of Mongolian descent, constitute around 9 percent of the population of Afghanistan and around 10 percent of the population of Balkh province (DHS, 2010)²⁸. They have historically been a marginalized group in Afghanistan with very different origins from the other ethnic groups in Afghanistan.²⁹ As stated earlier, although the two groups live in close proximity and they share the same language, their villages are fully ethnically segregated and there are very few economic interactions between the two areas.

Now I examine whether the main results differ by ethnicity. I do not make any predictions and take this as a purely empirical exercise. Table 1.A.8 shows that all the

²⁷A similar regression explaining seasonal changes in MAO by changes in average village-level income does not yield a significant result, though. Results are available upon request.

²⁸The remaining ethnic groups in Balkh province are Pashtuns (12 percent), Uzbeks (11 percent), Turkmen (9 percent), and Balochis (2 percent). The remaining 12 percent did not report their ethnicity. *Source:* Demographic and Health Survey Afghanistan (2010). Indian Institute for Health Management Research (IIHMR), available online at https://dhsprogram.com/data/dataset/Afghanistan_Special_2010.cfm.

²⁹Hazaras faced social, economic and political discrimination, often resulting in atrocities against members of the group. The massacres of Hazaras in 1880s during the reign of Abdur Rahman Khan, and later in 1994 in Kabul and in 1997 in Mazar-e-Sharif during the reign of the Taliban “irreparably damaged the fabric of the country’s national and religious soul” (Rashid 2001, p. 83). Hazaras were sidelined from mainstream Afghan politics when the 1964 constitution ruled that all state officials have to be Sunni (Hanafi) muslims. Although the new constitution does not continue to discriminate against Hazaras and there are many high ranking Hazara officials in the government, the ethnic division is still present.

main results are valid for both the Tajiks (Columns 1 to 3) as well as for the Hazaras (Columns 4 to 6) in my sample. That is, the transfers in both the DG and TPPG remain stable over time, and that the enforcement of sharing norms weakens substantially during the lean season.

1.5.4 Potential Confounds

The experiment was conducted over two periods, the lean season first and the post-harvest season second. What if the order of the experiments alone influences the results? Two findings refute such a claim.

First, it might be argued that the stability of sharing behavior I observe can be attributed to anchoring one's own behavior in the first, lean season experimental round. For this to be the case, the PAs would have to remember their behavior in the previous experimental round. When asked during the post-harvest round post-experimental survey—in an unincentivised question—about how much they transferred in the DG in the previous round, the PAs' guesses were correlated more with the actual transfers in the post-harvest round (0.61, $p < 0.01$), than with the transfers in the lean season round (0.48, $p < 0.01$). I asked this question only of PAs. Moreover, only about 32 percent of the participants (22 out of 68) correctly guessed their own transfer in the lean season round. Twelve of these 22 participants decided to choose the same amounts in both rounds. When conducting the same analysis as in Table 1.4 on a subsample of 46 PAs who did not remember their DG transfers from the previous round correctly, I obtain results that are qualitatively very similar to the results obtained for the full sample of 68 PAs, with no statistically significant differences in DG or TPPG transfers across seasons (see Table 1.A.9).

Second, to examine the possible role of order effects on punishment behavior, I compare the subjects who participated in both seasons and are thus susceptible to being influenced by the order of the experimental rounds to the "virgin" population of farmers participating in one season only—either in the lean season or in the post-harvest season. Reassuringly, the personal characteristics of farmers who participated in both seasons and those who participated in the lean season do not differ (Column 7 of Table 1.A.2), but the sample of participants recruited for the first time in the post-harvest period is significantly younger and less educated despite the same sampling procedure (Column 9 of Table 1.A.2). Table 1.A.10 shows that the punishment behavior of PCs who participated

in both periods is not statistically significantly different from the “virgin” subjects in the respective seasons (lean season: $F(1,182) = 0.65$, $p=0.42$; post-harvest season: $F(1,182) = 0.56$, $p=0.46$). Also, the difference between the sanctioning behavior of “virgin” PCs in the post-harvest season and in the lean season exhibits a very similar declining pattern as I observe among the participants in both periods in Table 1.5 ($F(1,182)=6.43$, $p=0.01$).

Another confound that might explain the results presented here is that of seasonal changes in marginal utility of wealth. It is plausible that, due to diminishing marginal utility of wealth, an additional ECU in the experiment has a different value in different seasons. This issue gains importance in the context of dramatic seasonal income fluctuations. As the marginal utility of an additional ECU is highest in the lean season on average, we should expect the participants to put a higher value on their own payoffs in the lean season, *ceteris paribus*. If that was the case, it would be possible to attribute the observed lower willingness to engage in punishment in the lean season to the diminishing marginal utility of wealth. I provide two arguments against this explanation.

First, it is not plausible that the changing marginal utility of wealth would result in a dramatic decline in punishment behavior, but not in the decline in sharing behavior during the lean season. This would imply a disproportionately lower elasticity of willingness to share with respect to wealth compared to the elasticity of willingness to punish with respect to wealth. Since willingness to share is positively correlated with willingness to punish (Henrich et al. 2006), such a conclusion is unlikely.³⁰

Second, as discussed in section 1.5.1, individual-level changes in income and poverty in general cannot explain the differences in behavior across seasons. It is thus rather inconceivable that changing marginal value of money across seasons is driving the observed behavioral change.³¹

1.6 Concluding Remarks

A large fraction of the world’s population is repeatedly exposed to periods of resource scarcity. Although there is a common understanding of social responses to extreme scarcities such as famines, when cooperation breaks down, we have much less of an understand-

³⁰My data on the village level also support a positive correlation between the willingness to share and the willingness to punish. Analysis available upon request.

³¹Similarly, if the participants were concerned about seasonal changes in marginal utility of wealth of their matched partners rather than their own, observation of the differential treatment in the sharing and in the sanctioning conditions would be equally unlikely.

ing of social responses to temporary periods of scarcity, common in many rural societies. In this paper I ask whether a society exposed to seasonal scarcity is able to sustain its informal sharing mechanisms. Specifically, I experimentally examine the dynamics of individual sharing behavior using a dictator game, and of the willingness of third parties to engage in the enforcement of sharing norms, using a third party punishment game among Afghan subsistence farmers. I visited the area twice in one year—during the lean season and six months later during a post-harvest season, the period of relative plenty—and conducted the same experiment repeatedly with the same participants.

Although the sharing behavior measured by the dictators' transfers in a standard dictator game remains stable over time on both the aggregate level, as well as, to a large extent, on the individual level, the enforcement mechanisms that help to sustain the cooperative outcomes—as measured by the intensity of third parties' willingness to punish non-desirable behavior—are significantly weakened during the period of scarcity.³² Even though the population studied seems to have developed some mechanisms to sustain prosociality over the period of temporary resource scarcities during the lean season, it is not implausible that cooperation might deteriorate if the population experiences a larger shock or if it is exposed to scarcity over a longer period of time than expected. This would be consistent with the decline in cooperation over time when enforcement mechanisms are not available, observed in previous laboratory experiments (Boyd et al. 2003; Fehr and Fischbacher 2004a). One can also speculate that with weakened norm enforcing behavior perverse behavior, such as anti-social punishment (Herrmann, Thöni, and Gächter 2008; Prediger, Vollan, and Herrmann 2014) or counter-punishment (Nikiforakis 2008), might gain prominence.

It is not clear how narrow this gap between cooperation and its breakdown is and more research should be done in this direction, but the present study offers some evidence that even temporary periods of resource scarcity substantially weaken the enforcement of sharing norms. Policymakers should take this finding seriously in addressing the issue of transitory scarcity, not only as a problem at the individual level, but also at the community level. More importantly, as mounting evidence on causal links between resource scarcity and the emergence of conflicts on a community level shows (Hsiang, Burke, and Miguel 2013) it is possible that many societies exposed to temporary periods of resource

³²Since sharing preferences are predictive of trusting and cooperative behavior, the results might have important implications for the functioning of markets and the ability of communities to mobilize and engage in collective action during periods of scarcity.

scarcity might be closer to a spark of violence than was previously thought. The herein observed erosion of social norm enforcement might be one of the explanatory factors.³³

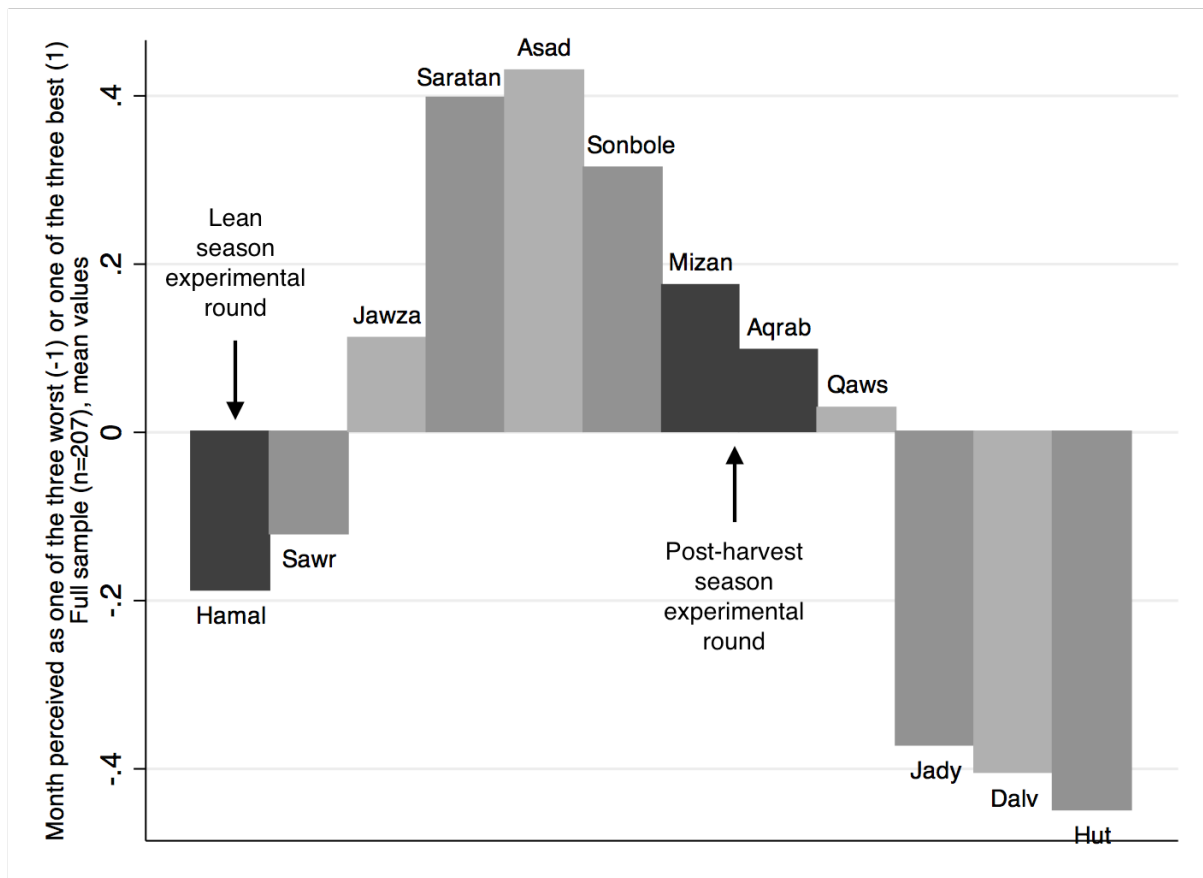
While focusing on the altruism and enforcement dimensions, my experiment abstracts from the possibility of informal risk-sharing based on reciprocal relationships, another crucial component of risk-sharing behavior. However, reciprocal relationships are predicted to be weaker in poorer populations where the marginal utility of current income is high and thus so is willingness to defect on relationships bringing potential gains in future (Coate and Ravallion 1993). This argument further weakens the risk-sharing opportunities during income shocks, such as seasonal scarcities. The prediction of the model needs to be taken with caution, as empirical evidence supporting it is missing, possibly due to endogeneity issues.

Policymakers already offer solutions to mitigate the seasonal scarcities and scarcities in general via the introduction of safety net programs (Alderman and Yemtsov 2014), the provision of formal insurance (Morduch 2006), the provision of microcredit (Banerjee 2013), and the introduction of reliable savings products (Dupas and Robinson 2013). While they usually promote the impact of these policies on individuals, they often fall short of stressing their possible effect on preventing negative outcomes on the community level. For example, an interesting unintended side-effect of a large-scale public employment program serving as a social insurance is that it reduces the risk of violent events (Fetzer 2014). Moreover, since scarcity is shown here to be associated with looser social norms enforcement, concerns that the introduction of such policies would crowd out existing informal institutions³⁴ seem less plausible.

³³See for example Sekhri and Storeygard (2014) or Blakeslee and Fishman (2013) who document an increase in violence and property crime as a response to rainfall failure in India or Oster (2004) and Miguel (2005) who document increased incidence of ritual murders after rainfall failures in renaissance Europe and in current rural Tanzania.

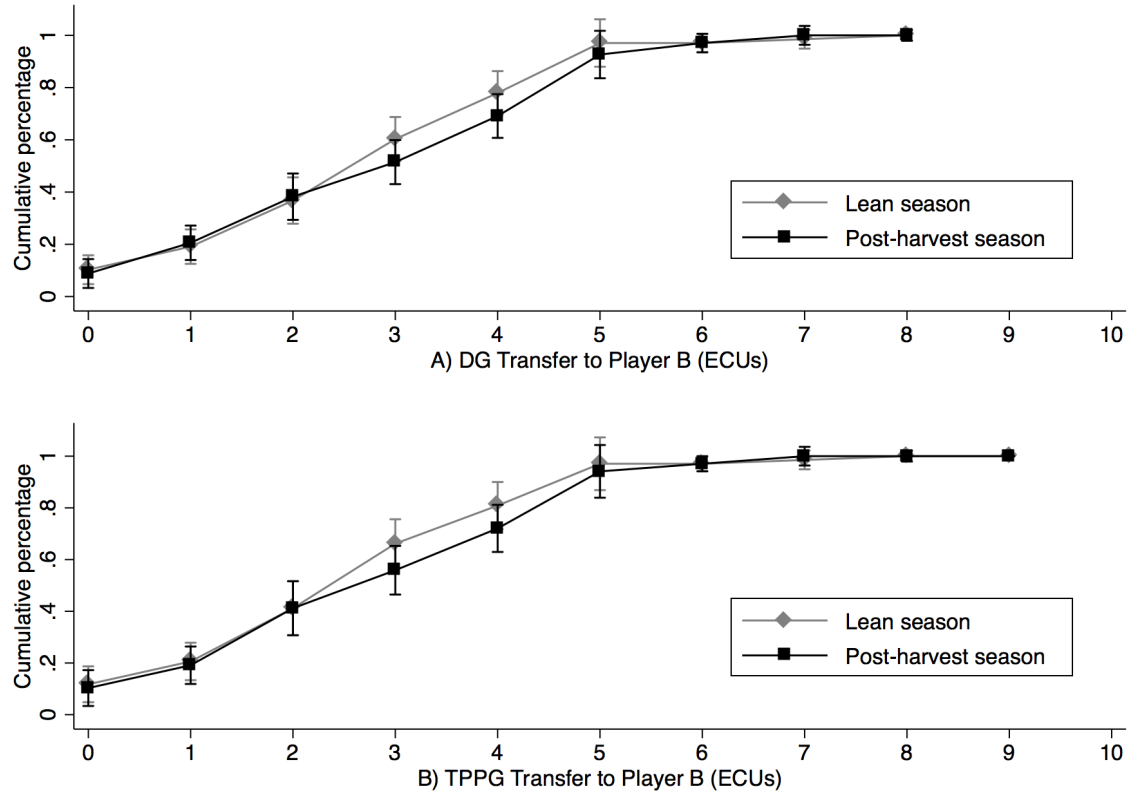
³⁴E.g., Dupas and Robinson (2013) observe that the introduction of a safe savings product results in decreased participation in existing informal insurance networks.

Figure 1.1: Subjective Perceptions of Living Quality Throughout the Year



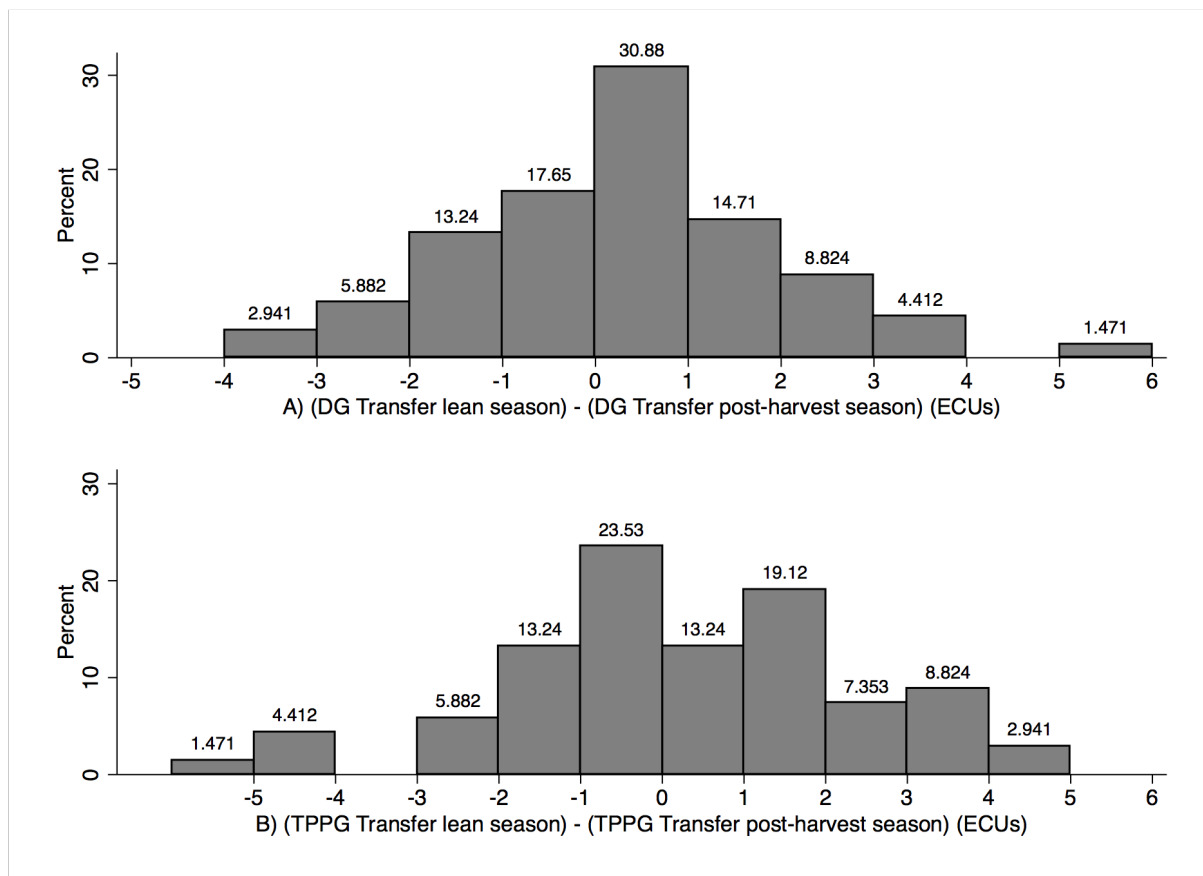
Notes: The figure depicts the average participants' rating of quality of life during each month in the year. The participants rated the month as one of the best three months (+1) or as one of the worst three months by answering the question: "Which three months are usually the [best /most difficult] in terms of food for you?". Months not mentioned are treated as 0. The question was asked during the lean season round. Afghanistan uses the Persian version of the Solar Hijri calendar. Persian month names are presented here, because the conversion to Gregorian calendar would be confusing. The experiments were carried out in the months of Hamal 1392 (March to April 2013, lean season) and Mizan and Aqrab 1392 (October 2013, post-harvest season) represented in the darkest color.

Figure 1.2: Cumulative Distributions of DG and TPPG Transfers Across Seasons



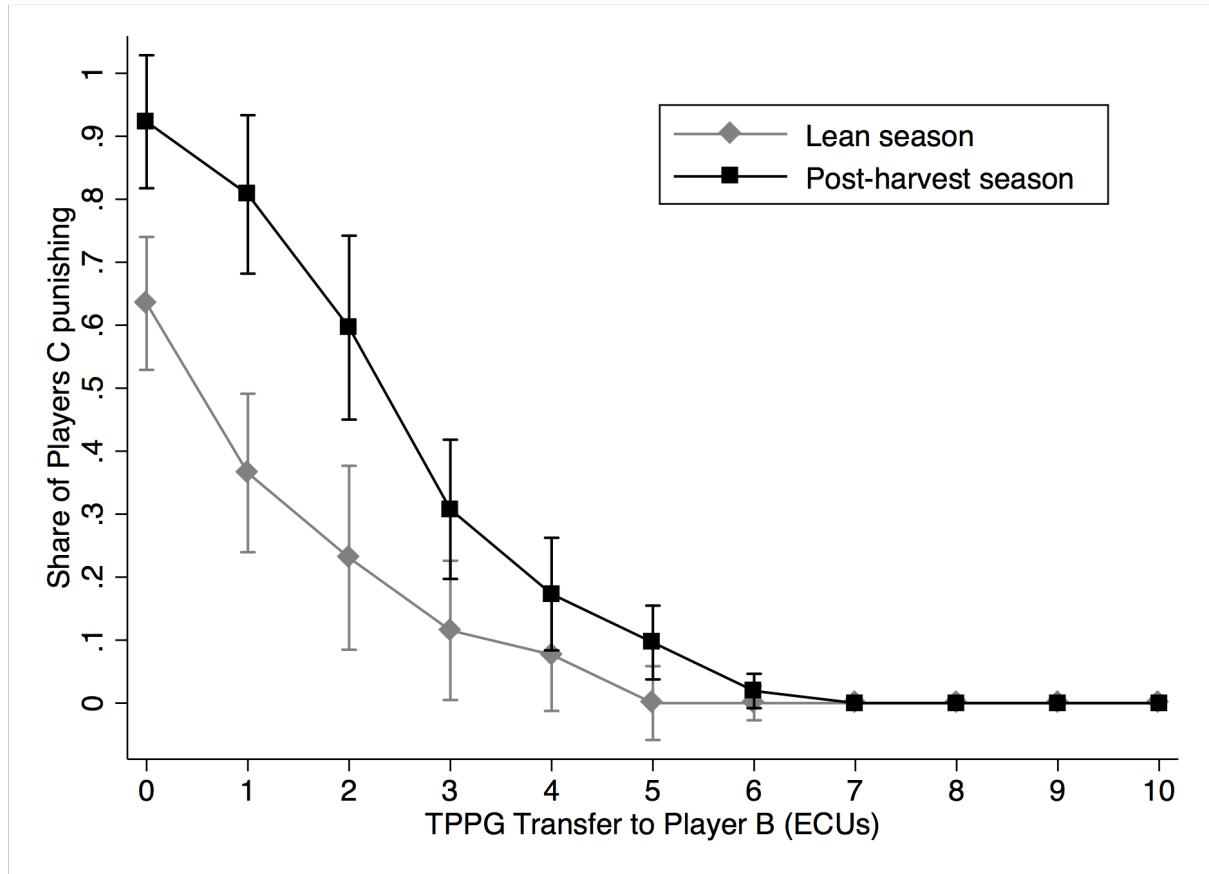
Notes: The figure shows the cumulative distribution of transfers from Player A (dictator) to Player B (passive receiver) in ECUs (allowed between 0 and 10) in A) the dictator game (DG) and B) the third party punishment game (TPPG) across the PAs participating in both rounds (n=68). The cumulative distribution of lean season transfers is depicted in grey, the cumulative distribution of post-harvest season transfers is depicted in black. The error bars represent 95 percent confidence intervals.

Figure 1.3: Distributions of Individual Changes in DG and TPPG Transfers Across Seasons



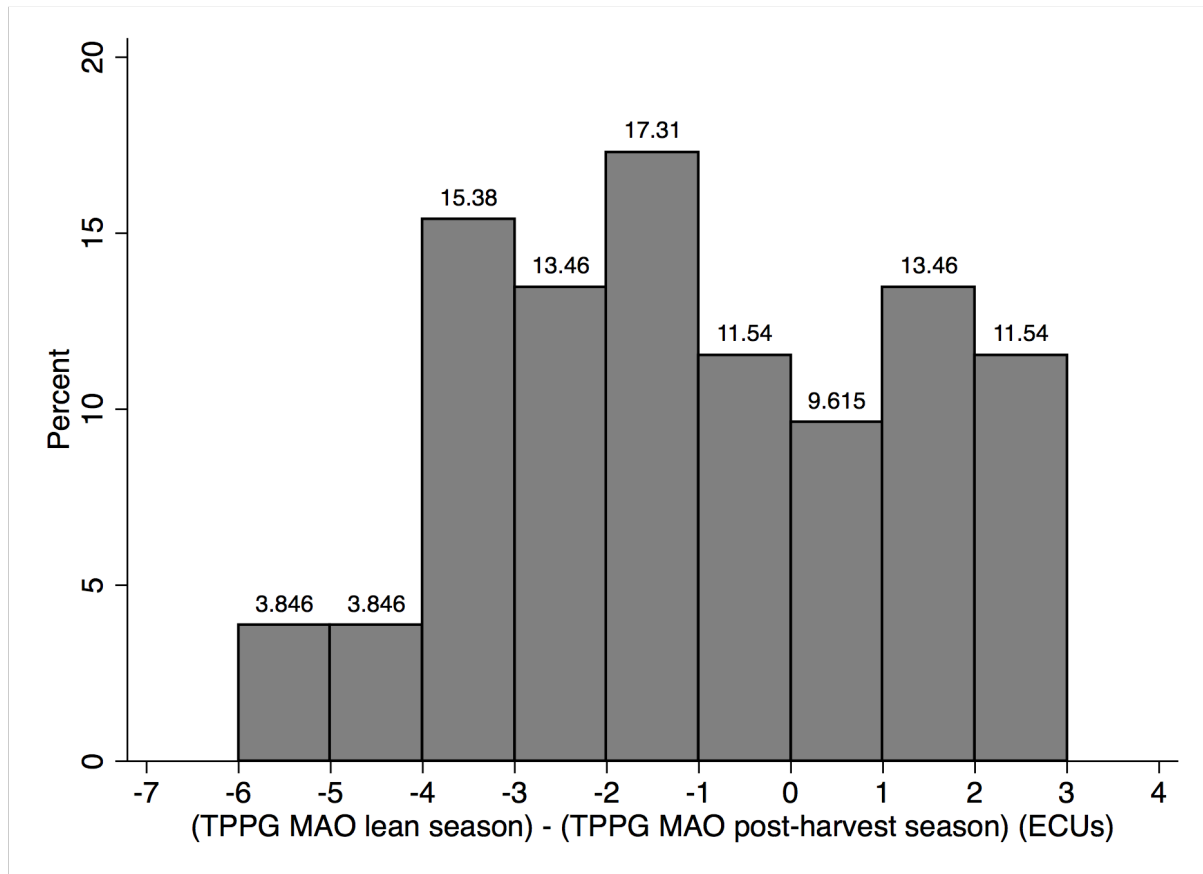
Notes: The figure shows the distributions of differences between the transfers in the lean season and the post-harvest season in A) the DG and B) the TPPG by participant. Transfer differences are in ECUs (the possible range is from -10 to 10).

Figure 1.4: Distributions of TPPG MAO Across Seasons



Notes: The figure shows the distribution of Player C (punisher; PC) minimum acceptable offers sent by Player A to Player B in the third party punishment game (TPPG MAO). I use data for the 52 PCs for whom TPPG MAO could be recovered in both rounds. The distribution of lean season TPPG MAO is depicted in grey, the distribution of post-harvest season TPPG MAO is depicted in black. The error bars represent 95 percent confidence intervals.

Figure 1.5: Distributions of Individual Changes in TPPG MAO Across Seasons



Notes: The figure shows the distribution of within-individual changes in Player C (punisher; PC) minimum acceptable offers sent by Player A to Player B in the third party punishment game (TPPG MAO) between the lean and the post-harvest season. I use data for the 52 PCs for whom TPPG MAO could be recovered in both rounds. Positive numbers represent higher TPPG MAO in the post-harvest season compared to the lean season.

Table 1.1: Descriptive Statistics

	Mean (1)	SD (2)
Age	38.83	(15.49)
Schooling (completed years)	2.97	(3.82)
Can read a letter (d)	0.58	(0.49)
Number of household members	9.66	(4.69)
Household head (d)	0.83	(0.38)
Not married (d)	0.11	(0.32)
Married to a single wife (d)	0.71	(0.45)
Married to multiple wives (d)	0.18	(0.38)
Daughters below 15 ^a	1.93	(1.66)
Sons below 15 ^a	2.13	(1.60)
Years living in village	36.98	(16.59)
Sunni (d)	0.51	(0.50)
Irrigated land (in jiribs)	4.47	(7.36)
Rainfed land (in jiribs)	10.81	(18.68)
Observations	207	

Notes: Means of the sample participating in both seasons are reported. Standard deviations in parentheses. ^a These questions were only asked to a sub-sample of A and C players (N=194).

Table 1.2: Seasonal Effects—Individual Time-Variant Characteristics

	Lean season		Post-harvest season		T-test	
	Mean (1)	SD (2)	Mean (3)	SD (4)	Difference (5)	t-value (6)
Cash earned in past 30 days (ths AFN) ^{a, b}	0.35	(0.79)	0.51	(0.62)	-0.16*	(-1.93)
Cash earned in past 30 days: selling food (ths AFN) ^{a, b}	0.15	(0.66)	0.31	(0.54)	-0.16*	(-2.18)
Cash earned in past 30 days: day labor (ths AFN) ^{a, b}	0.10	(0.26)	0.08	(0.28)	0.01	(0.38)
Perceived income situation ^b	-0.40	(0.67)	-0.03	(0.61)	-0.37***	(-5.89)
Meat eaten in past 7 days (times) ^a	0.73	(1.04)	0.98	(1.00)	-0.25*	(-2.05)
Now saves money (d) ^a	0.07	(0.26)	0.04	(0.20)	0.03	(1.02)
Now in debt (d) ^a	0.86	(0.34)	0.70	(0.46)	0.16***	(3.38)
Now providing loan (d) ^a	0.29	(0.45)	0.39	(0.49)	-0.10*	(-1.79)
Unable to work in past 30 days (days)	7.85	(10.09)	2.25	(6.83)	5.59***	(6.61)
Perceived stress score ^c	5.40	(1.99)	3.97	(1.15)	1.43***	(8.96)
Unusually high level of crop pests & diseases (d)	0.11	(0.32)	0.02	(0.14)	0.09***	(3.84)
Unusually high level of livestock diseases (d)	0.28	(0.45)	0.11	(0.32)	0.17***	(4.43)
Unusually high level of human disease (d)	0.50	(0.50)	0.20	(0.40)	0.30***	(6.70)
Participated in a dispute in past 30 days (d)	0.14	(0.35)	0.08	(0.27)	0.07**	(2.20)
Participated in a voluntary activity in past 30 days (d)	0.51	(0.50)	0.65	(0.48)	-0.14***	(-2.81)
Member of any village association now (d)	0.31	(0.46)	0.17	(0.38)	0.14***	(3.25)
Some household member migrated for work (d) ^a	0.25	(0.44)	0.24	(0.43)	0.01	(0.17)
Observations	207		207		414	

Notes: Means reported in Columns 1 and 3. Standard deviations in parentheses in Columns 2 and 4. Column 5 reports the difference between the means of respective characteristics in the post-harvest season and the lean season. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. Column 6 reports t-values of a two-sided t-test. ^a Questions asked of the subsample of N=139 Players A and C. ^b Cash earned by household head per OECD equivalence scaled household member. ^c Indicating whether the individual perceives his current income to be much worse (-2), worse (-1), same (0), better (+1), or much better (+2) relative to other fellow villagers. ^d A short version of the Cohen, Kamarck, and Mermelstein (1983) Perceived Stress Scale used: scale ranges from 0 to 8, 8 indicated the highest level of perceived stress.

Table 1.3: Seasonal Effects—Experimental Outcomes

	Lean season		Post-harvest season		T-test	
	Mean (1)	SD (2)	Mean (3)	SD (4)	Difference (5)	t-value (6)
<i>Player A (Dictator)</i>						
DG transfer (ECU)	3.03	(1.74)	3.22	(1.85)	-0.19	(-0.62)
TPPG transfer (ECU)	2.87	(1.74)	3.10	(1.82)	-0.24	(-0.77)
Belief: others' DG transfer (ECU)	2.94	(1.84)	3.04	(1.60)	-0.11	(-0.35)
Belief: others' TPPG transfer (ECU)	2.93	(1.63)	3.06	(1.67)	-0.13	(-0.44)
Belief: most PCs punish zero TPPG transfer (d)	0.72	(0.45)	0.71	(0.46)	0.01	(0.13)
Observations	68		68		136	
<i>Player B (Receiver)</i>						
Belief: others' DG transfer (ECU)	3.18	(2.03)	3.63	(1.61)	-0.46	(-1.45)
Belief: others' TPPG transfer (ECU)	3.66	(1.84)	3.68	(1.41)	-0.02	(-0.07)
Belief: most PCs punish zero TPPG transfer (d)	0.68	(0.47)	0.78	(0.42)	-0.10	(-1.35)
Observations	68		68		136	
<i>Player C (Punisher)</i>						
MAO (consistent responses; ECU) ^a	1.35	(1.51)	3.03	(1.87)	-1.68***	(-5.48)
Punish zero TPPG transfer (consistent responses; d) ^a	0.62	(0.49)	0.94	(0.25)	-0.32***	(-4.61)
Belief question about TPPG transfer (ECUs)	3.15	(1.71)	3.41	(1.56)	-0.26	(-0.91)
Belief: most PCs punish zero TPPG transfer (d)	0.65	(0.48)	0.79	(0.41)	-0.14*	(-1.88)
Observations	71		71		142	

Notes: Means reported in Columns 1 and 3. Standard deviations in parentheses in Columns 2 and 4. Column 5 reports the difference between the means of respective characteristics in the post-harvest season and the lean season. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. Column 6 reports t-values of a two-sided t-test. DG stands for the dictator game, TPPG stands for the third party punishment game, MAO stands for TPPG minimum acceptable offer. ^a Values reported for a subsample of N=123 observations (60 lean season, 63 post-harvest season) with consistent MAO.

Table 1.4: Effect of Seasonality on DG and TPPG Transfers

Dependent variable	DG transfer			TPPG transfer		
	(1)	(2)	(3)	(4)	(5)	(6)
Lean season	-0.19 (0.22)	-0.19 (0.22)	-0.11 (0.26)	-0.24 (0.27)	-0.24 (0.28)	-0.22 (0.35)
Age (in years/10)			-0.25* (0.15)			-0.19 (0.13)
Schooling (completed years)			-0.07 (0.06)			-0.06 (0.05)
Number of household members			-0.07 (0.05)			-0.05 (0.05)
Cash earned in past 30 days (ths AFN) ^a			-0.13 (0.14)			0.09 (0.18)
Poverty index (z-score)			-0.16 (0.18)			0.01 (0.20)
Village fixed effects	No	Yes	Yes	No	Yes	Yes
Constant	3.22*** (0.23)	3.38*** (0.88)	5.20*** (1.00)	3.10*** (0.22)	2.90*** (0.74)	4.26*** (0.85)
Observations	136	136	136	136	136	136
R-squared	0.00	0.17	0.26	0.00	0.14	0.19

Notes: OLS coefficients. Clustered standard errors in parentheses. Clustering at individual level. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. In Columns 1 to 3 the dependent variable is the dictator game (DG) transfer in ECUs (range from 0 to 10). In Columns 4 to 6 the dependent variable is the third party punishment game (TPPG) transfer in ECUs (range from 0 to 10). For one observation (id=5109) the poverty index measure was missing due to the fact that the participant did not respond to one of the survey questions. The missing observation was replaced by an average poverty index for the given round. The results are robust to replacement by a minimum as well as by a maximum poverty index amount (analysis available upon request). ^a Cash earned by household head per OECD equivalence scaled household member.

Table 1.5: Effect of Seasonality on TPPG MAO

Dependent variable	TPPG Minimum Acceptable Offer		
	(1)	(2)	(3)
Lean season	-1.68*** (0.31)	-1.68*** (0.32)	-1.49*** (0.34)
Age (in years/10)			-0.27** (0.13)
Schooling (completed years)			0.01 (0.06)
Number of household members			-0.06 (0.04)
Cash earned in past 30 days (ths AFN) ^a			-0.61** (0.28)
Poverty index (z-score)			-0.28* (0.15)
Village fixed effects	No	Yes	Yes
Constant	3.03*** (0.24)	3.43*** (0.54)	5.03*** (0.99)
Observations	123	123	123
R-squared	0.20	0.27	0.36

Notes: OLS coefficients. Clustered standard errors in parentheses. Clustering at individual level. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. The dependent variable in all models is the third party punishment game (TPPG) minimum acceptable offer (MAO). Subsample of N=123 observations (60 lean season, 63 post-harvest season) with consistent MAO. ^a Cash earned by household head per OECD equivalence scaled household member.

Table 1.6: Explaining Within-Individual Changes in MAO Across Seasons

Dependent variable	TPPG MAO Difference		
	(1)	(2)	(3)
Age (in years/10)	0.00 (0.29)	-0.00 (0.29)	-0.02 (0.29)
Schooling (completed years)	0.10 (0.10)	0.08 (0.09)	0.08 (0.11)
Number of household members	0.02 (0.10)	-0.01 (0.09)	-0.00 (0.10)
Cash earned in past 30 days (ths AFN) ^a - Lean season	-0.48 (0.55)	-0.36 (0.49)	
Cash earned in past 30 days (ths AFN) ^a - Post-harvest season	0.58 (1.15)		0.30 (1.05)
Poverty index (z-score) - Lean season	-0.21 (0.34)	-0.19 (0.32)	
Poverty index (z-score) - Post-harvest season	0.18 (0.55)		0.08 (0.53)
Village fixed effects	Yes	Yes	Yes
Constant	-3.59** (1.72)	-2.83 (1.69)	-2.52 (1.71)
Observations	52	52	52
R-squared	0.27	0.24	0.22

Notes: OLS coefficients. Robust standard errors in parentheses. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. The dependent variable in all models is the within-subject third party punishment game (TPPG) minimum acceptable offer (MAO) difference between MAO in the lean season and MAO in the post-harvest season. I control for village fixed effects in all models. Subsample of N=52 observations in each season with MAO consistent in both seasons. ^a Cash earned by household head per OECD equivalence scaled household member.

1.A Appendix 1

Table 1.A.1: Number of observations by village, role, including “virgin” subjects

Participating in...	Both seasons			Lean season only			Post-harvest season only		
	Player A	Player B	Player C	Player A	Player B	Player C	Player A	Player B	Player C
Abpartob	3	4	4	2	1	1	7	6	6
Baizai Bala	8	4	8	1	5	1	6	10	6
Jaw-Paya Ali Abad	4	7	6	6	3	4	10	7	8
Kalahkan Pain	8	8	6	2	2	4	7	7	9
Kalakhane-Bala	7	7	8	3	3	2	8	8	7
Kheirabad	3	2	2	2	3	3	7	8	8
Koche Aghaz	14	13	14	1	2	1	6	7	6
Marghzar	8	9	10	5	4	3	8	8	7
Quala-e-Noorak	8	7	8	2	3	2	7	8	7
Shuran-e-Bala	5	7	5	5	3	5	5	3	5
Total	68	68	71	29	29	26	71	72	69

Table 1.A.2: Descriptive Statistics Including the “Virgin” Subjects

	Both seasons		Lean season only		Post-harvest season only		T-test (1)-(3)		T-test (1)-(5)	
	Mean (1)	SD (2)	Mean (3)	SD (4)	Mean (5)	SD (6)	Difference (7)	t-value (8)	Difference (9)	t-value (10)
Age	38.83	(15.49)	37.25	(15.51)	33.50	(16.00)	-1.58	(-0.79)	-5.32***	(-3.43)
Schooling (completed years)	2.97	(3.82)	2.19	(3.16)	3.14	(4.14)	-0.78	(-1.65)	0.18	(0.45)
Can read a letter (d)	0.58	(0.49)	0.54	(0.50)	0.44	(0.50)	-0.04	(-0.66)	-0.14***	(-2.93)
Number of household members	9.66	(4.69)	9.20	(4.20)	8.60	(3.90)	-0.46	(-0.78)	-1.06**	(-2.50)
Household head (d)	0.83	(0.38)	0.77	(0.42)	0.61	(0.49)	-0.06	(-1.13)	-0.22***	(-5.19)
Not married (d)	0.11	(0.32)	0.13	(0.34)	0.33	(0.47)	0.02	(0.48)	0.23***	(5.77)
Married to a single wife (d)	0.71	(0.45)	0.69	(0.47)	0.61	(0.49)	-0.02	(-0.33)	-0.21***	(-4.82)
Married to multiple wives (d)	0.18	(0.38)	0.18	(0.39)	0.05	(0.22)	0.00	(0.00)	-0.02	(-0.99)
Daughters below 15 ^a	1.93	(1.66)	1.95	(1.39)	1.54	(1.51)	0.02	(0.07)	-0.2	(-1.04)
Sons below 15 ^a	2.13	(1.60)	1.93	(1.21)	1.82	(1.67)	-0.20	(-0.85)	0.03	(0.18)
Years living in village	36.98	(16.59)	34.95	(16.38)	32.01	(16.56)	-2.03	(-0.95)	4.25	(0.90)
Sunni (d)	0.51	(0.50)	0.51	(0.50)	0.49	(0.50)	0.00	(0.07)	-0.02	(-0.34)
Irrigated land (in jiribs)	4.47	(7.36)	3.58	(3.79)	3.74	(5.54)	-0.89	(-1.05)	-0.73	(-1.13)
Rainfed land (in jiribs)	10.81	(18.68)	9.67	(14.36)	9.76	(22.06)	-1.14	(-0.50)	-1.05	(-0.52)
Observations	207		84		204		291		411	

Notes: Means reported in Columns 1, 3, and 5. Standard deviations in parentheses in Columns 2, 4, and 6. Column 7 reports the difference between the means of the respective characteristics for the sample of participants in both seasons and the sample of participants in the lean season only. Column 9 reports the difference between the means of the respective characteristics for the sample of participants in both seasons and for the sample of participants in the post-harvest season only. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. Columns 8 and 10 report t-values of a two-sided t-test. ^a Questions asked to the subsample of N=139 Players A and C in both periods, N=55 in the lean season, and N=136 in the post-harvest season.

Table 1.A.3: Village Level Effects

Dependent variable	DG transfer			TPPG transfer		
	Full (1)	Lean (2)	Post-harvest (3)	Full (4)	Lean (5)	Post-harvest (6)
Abpartob	1.21 (0.81)	1.19 (1.21)	1.24 (1.15)	1.55* (0.85)	1.76 (1.20)	1.33 (1.31)
Baizai Bala	0.40 (0.77)	0.36 (1.14)	0.45 (1.11)	0.90 (0.73)	1.05 (1.11)	0.75 (1.04)
Jaw-Paya Ali Abad	-0.29 (0.77)	-0.39 (1.15)	-0.18 (1.10)	0.34 (0.86)	-0.82 (1.20)	1.50* (0.80)
Kalahkan Pain	-0.41 (0.81)	-0.27 (1.20)	-0.55 (1.18)	-0.41 (0.75)	-0.32 (1.20)	-0.50 (0.99)
Kheirabad	1.05 (0.94)	-0.14 (1.18)	2.24** (1.03)	1.38 (0.98)	0.10 (1.20)	2.67*** (0.96)
Koche Aghaz	-1.32* (0.71)	-1.36 (1.14)	-1.29 (0.93)	-0.64 (0.71)	-0.71 (1.13)	-0.57 (0.94)
Marghzar	0.40 (0.78)	1.11 (1.20)	-0.30 (1.06)	0.59 (0.76)	1.43 (1.22)	-0.25 (0.95)
Quala-e-Noorak	0.09 (0.76)	0.23 (1.17)	-0.05 (1.05)	0.34 (0.70)	0.80 (1.10)	-0.12 (0.94)
Shuran-e-Bala	-0.39 (0.82)	-0.34 (1.42)	-0.43 (0.95)	0.21 (0.76)	0.83 (1.31)	-0.40 (0.85)
Constant	3.29*** (0.65)	3.14*** (1.06)	3.43*** (0.84)	2.79*** (0.62)	2.57** (1.05)	3.00*** (0.76)
Observations	136	68	68	136	68	68
R-squared	0.16	0.21	0.20	0.13	0.23	0.21

Notes: OLS coefficients. The constant represents the omitted village, Kalakhan-e-Bala. Robust standard errors in parentheses. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. In Columns 1 to 3 the dependent variable is the dictator game (DG) transfer in ECUs (range from 0 to 10). In Columns 4 to 6 the dependent variable is the third party punishment game (TPPG) transfer in ECUs (range from 0 to 10).

Table 1.A.4: Effect of Seasonality on DG and TPPG Transfers (Ordered Probit)

Dependent variable	DG transfer of...						TPPG transfer...					
	... 0	... 1	... 2	... 3	... 4	... 5	... 0	... 1	... 2	... 3	... 4	... 5
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Lean season	0.01 (0.02)	0.01 (0.02)	0.01 (0.02)	0.00 (0.00)	-0.01 (0.02)	-0.02 (0.04)	0.02 (0.03)	0.01 (0.02)	0.02 (0.03)	-0.00 (0.00)	-0.01 (0.02)	-0.03 (0.05)
Age (in years/10)	0.02 (0.01)	0.02 (0.01)	0.02* (0.01)	0.00 (0.00)	-0.01 (0.01)	-0.04* (0.02)	0.02 (0.01)	0.01 (0.01)	0.02 (0.01)	-0.00 (0.00)	-0.01 (0.01)	-0.03 (0.02)
Schooling (completed years)	0.01 (0.00)	0.01 (0.00)	0.01 (0.00)	0.00 (0.00)	-0.00 (0.00)	-0.01 (0.01)	0.01 (0.00)	0.00 (0.00)	0.01 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.01 (0.01)
Number of household members	0.01 (0.00)	0.01 (0.00)	0.01 (0.00)	0.00 (0.00)	-0.00* (0.00)	-0.01 (0.01)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.01 (0.01)
Cash earned in past 30 days (ths AFN) ^a	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.00 (0.00)	-0.01 (0.01)	-0.02 (0.02)	-0.01 (0.02)	-0.01 (0.01)	-0.01 (0.01)	0.00 (0.00)	0.01 (0.01)	0.01 (0.02)
Poverty index (z-score)	0.01 (0.01)	0.01 (0.01)	0.02 (0.02)	0.00 (0.00)	-0.01 (0.01)	-0.03 (0.03)	0.00 (0.02)	0.00 (0.01)	0.00 (0.02)	-0.00 (0.00)	-0.00 (0.01)	-0.00 (0.03)
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	136	136	136	136	136	136	136	136	136	136	136	136

Notes: Ordered probit. Average marginal effects on the probability of respective DG (columns 1-6) and TPPG (columns 7-12) transfers reported. Excluding marginal effects for infrequent transfers over 5 ECU. Clustered standard errors in parentheses. Clustering at individual level. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. In Columns 1 to 3 the dependent variable is the dictator game (DG) transfer in ECUs (range from 0 to 10). In Columns 4 to 6 the dependent variable is the third party punishment game (TPPG) transfer in ECUs (range from 0 to 10). For one observation the poverty index measure was missing due to the fact that the participant did not respond to one of the survey questions. The missing observation was replaced by an average poverty index for the given round. The results are robust to replacing by a minimum as well as by a maximum poverty index amount (analysis available upon request). ^a Cash earned by household head per OECD equivalence scaled household member.

Table 1.A.5: Effect of Seasonality on TPPG MAO (Ordered Probit)

Dependent variable	TPPG Minimum Acceptable Offer of...					
	... 0 (1)	... 1 (2)	... 2 (3)	... 3 (4)	... 4 (5)	... 5 (6)
Lean season	0.27*** (0.07)	0.14*** (0.04)	-0.01 (0.02)	-0.13*** (0.04)	-0.11*** (0.03)	-0.10*** (0.03)
Age (in years/10)	0.05** (0.02)	0.03** (0.01)	-0.00 (0.00)	-0.03** (0.01)	-0.02** (0.01)	-0.02* (0.01)
Schooling (completed years)	-0.00 (0.01)	-0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Number of household members	0.01 (0.01)	0.01* (0.00)	-0.00 (0.00)	-0.01* (0.00)	-0.01 (0.00)	-0.00 (0.00)
Cash earned in past 30 days (ths AFN) ^a	0.13** (0.05)	0.07** (0.04)	-0.00 (0.01)	-0.07* (0.04)	-0.06** (0.02)	-0.05* (0.03)
Poverty index (z-score)	0.05* (0.03)	0.03* (0.02)	-0.00 (0.00)	-0.03 (0.02)	-0.02* (0.01)	-0.02* (0.01)
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	123	123	123	123	123	123

Notes: Ordered probit. Average marginal effects on the probability of respective TPPG MAO reported. Excluding marginal effects for infrequent TPPG MAO over 5. Clustered standard errors in parentheses. Clustering at individual level. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. The dependent variable in all models is the third party punishment game (TPPG) minimum acceptable offer (MAO). Subsample of N=123 observations (60 lean season, 63 post-harvest season) with consistent MAO. ^a Cash earned by household head per OECD equivalence scaled household member.

Table 1.A.6: Average Village-Level Changes in TPPG MAO and in Disputes Engagement

Dependent variable	Village average of			Individual		
	TPPG MAO _{lean} -TPPG MAO _{post-harvest}			TPPG MAO _{lean} -TPPG MAO _{post-harvest}		
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Village-level "Engaged in disputes (d)" ^a	-7.93*** (2.14)	-7.98*** (2.15)	-7.51** (2.46)	-6.69* (3.63)	-6.83* (3.78)	-6.47* (3.61)
Constant	-0.84* (0.44)	-0.80 (0.44)	-0.98 (0.57)	-0.98** (0.37)	-0.95** (0.38)	-1.10*** (0.39)
Observations	10	10	10	52	52	52
R-squared	0.44	0.41	0.34	0.07	0.07	0.06
Weight used	No weight	Sample population	Village population	No weight	Sample population	Village population

Notes: OLS coefficients. Columns 2, 3, 5, and 6 report weighted data using analytic weights. Weights used are the sample population and the reported population of the entire village based on interviews with community leaders for Columns 2 and 5, and 3 and 6 respectively. Robust standard errors in parentheses. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. The dependent variable in models 1 to 3 is the difference in village-level average Third Party Punishment Game (TPPG) Minimum Acceptable Offer (MAO) in the lean season minus the post-harvest season TPPG MAO. The dependent variable in models 4 to 6 is the difference in individual Third Party Punishment Game (TPPG) Minimum Acceptable Offer (MAO) in the lean season minus the post-harvest season TPPG MAO. Observations in models 1 to 3 represent villages. ^a The lean season minus post-harvest season change in average village level engagement of individuals in disputes. Individuals asked if they "participated in a dispute in the previous four weeks".

Table 1.A.7: Average Changes in TPPG MAO and Intensity of Scarcity

Dependent variable	Village average of			Individual		
	TPPG MAO _{lean} -TPPG MAO _{post-harvest}			TPPG MAO _{lean} -TPPG MAO _{post-harvest}		
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Poverty z-score ^a	-2.87* (1.35)	-2.56* (1.35)	-3.58** (1.30)	-2.07* (1.09)	-1.90* (1.11)	-2.75** (1.10)
Constant	-1.34*** (0.35)	-1.33*** (0.36)	-1.25*** (0.33)	-1.52*** (0.32)	-1.52*** (0.31)	-1.46*** (0.32)
Observations	10	10	10	52	52	52
R-squared	0.33	0.31	0.45	0.06	0.05	0.10
Weight used	No weight	Sample population	Village population	No weight	Sample population	Village population

Notes: OLS coefficients. Columns 2, 3, 5, and 6 report weighted data using analytic weights. Weights used are the sample population and the reported population of the entire village based on interviews with community leaders for Columns 2 and 5, and 3 and 6 respectively. Robust standard errors in parentheses. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. The dependent variable in models 1 to 3 is the difference in village-level average Third Party Punishment Game (TPPG) Minimum Acceptable Offer (MAO) in the lean season minus the post-harvest season TPPG MAO. The dependent variable in models 4 to 6 is the difference in individual Third Party Punishment Game (TPPG) Minimum Acceptable Offer (MAO) in the lean season minus the post-harvest season TPPG MAO. Observations in models 1 to 3 represent villages. ^a The lean season minus post-harvest season change in average village level engagement of individuals in disputes. Individuals asked if they "participated in a dispute in the previous four weeks".

Table 1.A.8: Effect of Seasonality on DG Transfers, TPPG Transfers, and TPPG MAO (by Ethnic Group)

Dependent variable	Tajik			Hazara		
	DG transfer (1)	TPPG transfer (2)	TPPG Minimum Acceptable Offer (3)	DG transfer (4)	TPPG transfer (5)	TPPG Minimum Acceptable Offer (6)
Lean season	-0.02 (0.37)	0.39 (0.38)	-2.11*** (0.53)	-0.17 (0.41)	-0.72 (0.59)	-1.20** (0.49)
Age (in years/10)	-0.22 (0.19)	-0.09 (0.16)	-0.47*** (0.16)	-0.43** (0.16)	-0.44** (0.18)	-0.02 (0.22)
Schooling (completed years)	-0.16* (0.08)	-0.10 (0.07)	-0.04 (0.05)	0.04 (0.08)	-0.01 (0.07)	0.03 (0.10)
Number of household members	-0.09 (0.07)	-0.07 (0.06)	-0.08** (0.04)	-0.01 (0.05)	0.01 (0.06)	-0.09 (0.09)
Cash earned in past 30 days (ths AFN) ^a	0.05 (0.41)	0.79** (0.36)	-1.79*** (0.55)	-0.12 (0.15)	0.02 (0.16)	-0.43 (0.31)
Poverty index (z-score)	-0.16 (0.23)	-0.09 (0.18)	-0.21 (0.22)	-0.14 (0.28)	0.28 (0.36)	-0.45** (0.18)
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	5.57*** (1.10)	3.68*** (0.93)	6.95*** (1.34)	7.36*** (0.90)	6.98*** (0.88)	4.94*** (0.93)
Observations	72	72	63	64	64	60
R-squared	0.24	0.23	0.50	0.41	0.27	0.31
$\beta_{Lean\ season, Tajik} - \beta_{Lean\ season, Hazara}$	0.14	1.11	-0.91			
F-test p-values	(0.78)	(0.08)	(0.16)			

Notes: OLS coefficients. Clustered standard errors in parentheses. Clustering at individual level. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. In Columns 1 and 4 the dependent variable is the dictator game (DG) transfer in ECUs (range from 0 to 10). In Columns 2 to 5 the dependent variable is the third party punishment game (TPPG) transfer in ECUs (range from 0 to 10). In Columns 3 to 6 the dependent variable in all models is the third party punishment game (TPPG) minimum acceptable offer (MAO). The last two rows compare the coefficients on *Lean season* from both *Tajik* and *Hazara* regressions, using an F-test. ^a Cash earned by household head per OECD equivalence scaled household member.

Table 1.A.9: Effect of Seasonality on DG and TPPG Transfers (Subsample of PAs Who Do Not Recall Their Own Previous Round DG Transfer)

Dependent variable	DG transfer			TPPG transfer		
	(1)	(2)	(3)	(4)	(5)	(6)
Lean season	-0.13 (0.29)	-0.13 (0.31)	-0.04 (0.40)	-0.15 (0.35)	-0.15 (0.37)	-0.05 (0.48)
Age			-0.03 (0.02)			-0.02 (0.01)
Schooling (completed years)			-0.10 (0.08)			-0.09 (0.06)
Number of household members			-0.08 (0.07)			-0.05 (0.07)
Cash earned in past 30 days (ths AFN)			-0.08 (0.06)			0.02 (0.07)
Poverty index (z-score)			-0.23 (0.20)			-0.07 (0.20)
Village fixed effects	No	Yes	Yes	No	Yes	Yes
Constant	3.28*** (0.28)	3.87*** (1.13)	5.96*** (1.40)	3.15*** (0.26)	3.48*** (0.82)	4.93*** (1.22)
Observations	92	92	92	92	92	92
R-squared	0.00	0.18	0.28	0.00	0.12	0.18

Notes: OLS coefficients. Clustered standard errors in parentheses. Clustering at individual level. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. In Columns 1 to 3 the dependent variable is the dictator game (DG) transfer in ECUs (range from 0 to 10). In Columns 4 to 6 the dependent variable is the third party punishment game (TPPG) transfer in ECUs (range from 0 to 10). Subsample of 46 PAs who did not recall their DG transfers from the previous, lean season round.

Table 1.A.10: Differences TPPG MAO by Subjects Participating in Both Rounds and in One Round Only

Dependent variable	TPPG Minimum Acceptable Offer		
	(1)	(2)	(3)
Lean season ("virgin")	1.87*** (0.35)	1.71*** (0.52)	2.72*** (0.81)
Lean season (both seasons)	1.35*** (0.19)	1.41*** (0.39)	2.36*** (0.70)
Post-harvest season (both seasons)	3.03*** (0.24)	3.07*** (0.44)	3.72*** (0.70)
Post-harvest season ("virgin")	3.48*** (0.29)	3.38*** (0.42)	3.95*** (0.62)
Age (in years/10)			-0.07 (0.10)
Schooling (completed years)			0.04 (0.04)
Number of household members			-0.05 (0.03)
Cash earned in past 30 days (ths AFN) ^a			-0.45 (0.30)
Poverty index (z-score)			-0.37*** (0.13)
Village fixed effects	No	Yes	Yes
Observations	200	200	200
R-squared	0.68	0.71	0.73

Notes: OLS coefficients. Regression without a constant. Robust standard errors in parentheses. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. The dependent variable in all models is the third party punishment game (TPPG) minimum acceptable offer (MAO). Subsample of N=200 observations (23 lean season "virgin", 60 lean season participating in both seasons, 63 post-harvest season participating in both seasons, and 57 post-harvest season "virgin") with consistent MAO. ^a Cash earned by household head per OECD equivalence scaled household member.

1.A.1 Image Documentation

Figure 1.A.1: Explaining Instructions in a Group



(a) Experimental Subjects

(b) Explaining Instructions in a Group

Figure 1.A.2: Individual Player Experimental Sessions



1.A.2 Experiment Instructions

Group General Instructions

Before we begin I want to tell you about what we are doing here today and explain the rules that we must follow. We will be making a task in which you can get some money. Whatever money you will get in the task will be yours to keep and take home.

Maybe you won't get any money from the task, but if you decide to stay with us today, I will pass out 100 AFN to each of you to thank you for coming today. This money is not part of the task, it will be yours to keep. You will also get some snack and tea when you finish the task.

You should understand that this is not our own money. A University gave this money to us for research. This payment will not be regularly repeated in the future. It is not assistance, you will get the money for the task you will do here for us. It is not even a survey that you may have experienced before.

Please, also understand that there is no relation between our University and the organization People in Need delivering assistance in this area for a long period. I will not tell the organization about what you did here. Also, nothing you do here today will affect how the organization treats you or your community.

You should understand that there are no "right" or "wrong" answers in this task. Also, let me stress something that is very important. You were invited here without understanding what we are planning to do today. If you find that this is something that you do not wish to participate in, you can leave anytime.

Now, I will explain the task to you in the group. Later one after the other will come with me to carry out the task. It is important that you listen as carefully as possible, because only people who understand the task will actually be invited to participate. We will run through some examples here while we are all together.

You cannot ask questions or talk while we are here in the group. This is very important. Please be sure that you obey this rule, because it is possible for one person to spoil the task for everyone. If one person talks about the task while sitting in the group, we will not be able to carry out the task today. But do not worry if you do not completely understand the task as I show you the examples here in the group. Each of you will have time to ask questions when we sit alone together to be sure that you understand what you have to do. Now I will explain you what we are going to do during the task.

Group Games Instructions: Dictator Game

In one part of the task there will be two persons - Person A, and Person B. Both persons come from this village. None of you will know exactly with whom you are interacting. Only I know who will interact with whom and I will never tell anyone else.

Here are 200 AFN in 20 AFN bills that I will give to a Person A. Person A must decide how much of these 200 AFN he wants to give to Person B and how much he wants to keep for himself. I will not give any money to Person B. Person B takes home whatever Person A gives to him.

Here are some examples:

1. Suppose Person A gives 100 AFN to Person B, and keeps 100 AFN for himself. Person A goes home with 100 AFN (From the 200 AFN he had given 100 AFN to Person B and had kept 100

AFN for himself). Person B goes home with the 100 AFN from Person A.

2. Here is another example. Suppose Person A gives 0 AFN to Person B and keeps 200 AFN for himself. In this case, Person A goes home with 200 AFN. Person B doesn't have anything.
3. Here is another example. Suppose Person A gives 200 AFN to Person B and keeps 0 AFN for himself. In this case, Person A goes home with 0 AFN. Person B goes home with the 200 AFN from Person A.
4. Here is another example. This time suppose Person A gives 60 AFN to Person B and keeps 140 AFN for himself. In this case, Person A goes home with 140 AFN. Person B goes home with the 60 AFN from Person A.

Note again, there are no "right" or "wrong" answers in this task.

Group Games Instructions: Third Party Punishment Game

In another part of the task, there will be three persons - Person A, Person B, and Person C. All three persons come from this village. None of you will know exactly with whom you are interacting, but it will definitely not be the person with which you interacted in the previous part of the task. Only I know who will interact with whom and I will never tell anyone else.

Here is another 200 AFN. Person A must decide how much of these 200 AFN he wants to give to Person B and how much he wants to keep for himself. Person B takes home whatever Person A gives to him, but Person A has to wait until Person C has made a decision before finding out what he is going to take home. Person C is given 100 AFN. Person C can make three things with his 100 AFN.

1. He can pay 20 AFN to subtract 60 AFN of Person A's money, which Person A wanted to keep for himself. This money will be taken away; none of the Persons will get it. Person C will keep the remaining 80 AFN.
2. He can pay 40 AFN to subtract 120 AFN of Person A's money, which Person A wanted to keep for himself. This money will be taken away; none of the Persons will get it. Person C will keep the remaining 60 AFN.
3. He can pay nothing, keep all of the 100 AFN for himself and leave the money Person A wanted to keep for himself untouched.

Before hearing how much Person A has given to Person B, Person C has to decide what he wants to do for each of the possible amounts that Person A can give to Person B. This is 0 AFN, 20 AFN, 40 AFN, 60 AFN, 80 AFN, 100 AFN, 120 AFN, 140 AFN, 160 AFN, 180 AFN, or 200 AFN.

Here are some examples (All examples are shown with 20 AFN banknotes):

1. Here is another example. Suppose Person A gives 200 AFN to Person B and keeps 0 AFN for himself. Person C states that he would "do nothing" if Person A does this. In this case, Person A goes home with 0 AFN. Person B goes home with the 200 AFN from Person A, and Person C goes home with 100 AFN.

2. Here is another example. Suppose Person A gives 60 AFN to Person B and keeps 140 AFN for himself. Person C states that he would “do nothing” if Person A does this. In this case, Person A goes home with 140 AFN (He had kept 140 AFN for himself and Person C didn’t decide to subtract money from him). Person B goes home with the 60 AFN from Person A. And Person C goes home with 100 AFN.
3. Here is another example. As before, Person A gives 60 AFN to Person B and keeps 140 AFN for himself. But now, Person C states that he would pay 20 AFN to subtract 60 AFN from Person A’s money. In this case, Person A goes home with 80 AFN (He had kept 140 AFN for himself minus the 60 AFN equals 80 AFN). Person B goes home with the 60 AFN from Person A. And Person C goes home with 80 AFN.
4. And a last example: Suppose Person A gives 120 AFN to Person B and keeps 80 AFN for himself. Person C states that he would pay 20 AFN to subtract 60 AFN from Person A’s money. In this case, Person A goes home with 20 AFN (He had kept 80 AFN for himself minus the 60 AFN equals 20 AFN). Person B goes home with the 120 AFN from Person A. And Person C goes home with 80 AFN (100 AFN minus 20 AFN equals 80 AFN).

Again, there are no “right” or “wrong” answers in this task.

We will then call each of you in turn to make the task, starting with the person who picked number 1. In case you cannot read numbers, we will assist you.

When you finish the task, you have to wait until everybody has finished. Then I will call you in one by one again and I will tell you whether you have gained something. If yes, I will pay you that amount plus you will get the 100 AFN I promised you at the beginning.

We will not pay you for both tasks. At the end of the session you will have to pick a ball from a pouch to decide for which of the tasks you will get the payment. We will then give you the payment according to what color of the ball you picked. Please, take both tasks as if there was no other task before or after. Do you understand this?

Remember that you are not allowed to talk to the people still waiting to carry out the task. If you do talk to other people, the Assistant 3 will tell you to leave and not come back even if you may have earned some money.

Chapter 2

Contract Enforcement and Trustworthiness Across Ethnic Groups: Experimental Evidence from Northern Afghanistan

Vojtěch Bartoš and Ian Levely¹

Abstract

We study how the availability and use of a specific formal institution – a financial sanction – affects trust, trustworthiness, and moral intentions towards co-ethnics and non-co-ethnics using an economic experiment run with 420 adult males from peri-urban areas in Afghanistan. In contrast to previous studies on the behavioral effects of financial incentives, our subjects have little experience with formal institutions. We use a trust game with a requested back-transfer in which the investor can choose to impose a financial sanction for non-compliance. The sanction is costly to the trustee but cost-less to the investor. While sanctioning increases back-transfers in cross-ethnic pairs, it does not in co-ethnic pairs. Our results suggest that financial sanctions may crowd out moral incentives more strongly among one's own group, but have a much smaller behavioral effect when applied to individuals from a different ethnic group. The results have important implications for understanding how formal institutions affect cooperation in ethnically

¹We thank Abigail Barr, Michal Bauer, Subhasish M. Chowdhury, Guillaume Frechette, Peter Katuščák, Klára Kalíšková, Pieter Serneels and the participants in seminars at CERGE-EI, NYU, Rutgers University, and the University of East Anglia for their helpful comments, and Ahmad Qais Daneshjo and Hadia Essazada, for their excellent research assistance. This research was supported by a grant from the CERGE-EI Foundation under a program of the Global Development Network (RRC13+11), the Grant Agency of Charles University (46813), and the Czech Science Foundation (13-20217S).

heterogeneous settings.

2.1 Introduction

While formal contracts are essential to the functioning of modern, large-scale societies, because of costs and other difficulties associated with enforcement, trust plays an essential role in virtually all economic transactions (Arrow 1972; Tirole 2011). The average level of interpersonal trust in a country has been shown to predict economic performance (Knack and Keefer 1997), and when formal institutions for enforcing contracts are weak or absent, as is the case in many developing economies, trust and trustworthiness play an even greater role. However, in interactions with members of different ethnic and social groups, interpersonal trust may be lower (Alesina and La Ferrara 2005), which discourages trade and other cooperative interactions with individuals from more distant social groups.²

Because the creation and maintenance of efficient, formal institutions requires a high capacity for collective action, ethnically heterogeneous societies that have lower levels of trust—and thus might benefit the most from formal institutions that complement trust-based interactions—may also be the least likely to develop them (Collier 1999). Given this, facilitating the emergence of formal institutions, and strengthening existing informal arrangements are top priorities of governments and development agencies, which promote the introduction of formal governance bodies in many developing countries (Shah 2006).³

Yet a number of studies have shown that formal and informal institutions do not always complement one another; the introduction of material incentives—such as the introduction of enforceable contracts—can crowd out “moral incentives” for adhering to informal norms guiding mutual cooperation (Bowles 2008; Frey and Jegen 2001; Bowles

²Alesina and La Ferrara (2005) find that more ethnically heterogeneous countries exhibit lower levels of generalized trust and are also less wealthy on average. Miguel and Gugerty (2005) find that more ethnically heterogeneous communities in western Kenya are less likely to invest in schools or communal water wells, and similar behavior has been studied in American cities (Alesina and La Ferrara 2002; Costa and Kahn 2003). Habyarimana et al. (2007) show experimentally that ethnically heterogeneous groups in Uganda are less able to cooperate in a public goods environment, due to failure to collaborate in enforcing norms for collective action. Ingroup favoritism in trust and trustworthiness is a phenomenon experimentally documented both in developing countries (Johansson-Stenman, Mahmud, and Martinsson 2009) and in developed countries (Falk and Zehnder 2013).

³In Afghanistan specifically, the World Bank and the national government are currently implementing a nation-wide National Solidarity Program aimed to introduce formal local governance bodies through which small infrastructure development grants are channeled. Between 2003 and 2013, almost 31,000 Community Development Councils were introduced nationwide.

and Polania-Reyes 2012).⁴ A better understanding of whether extrinsic incentives complement or substitute for an individual’s intrinsic motivations would provide a deeper understanding of human behavior and have implications for policy makers in settings with weak formal institutions. In this study we examine how the introduction of formal sanctions affects trust and trustworthiness, both within and across ethnic groups. We conducted an economic experiment in peri-urban areas of Mazar-e-Sharif in northern Afghanistan, a country currently undergoing an exogenously coordinated transition from informal to formal institutions, and where ethnic identity plays an important role in everyday life. The post-conflict state-building efforts in an ethnically heterogeneous environment make Afghanistan an ideal setting for our case. The subjects in our experiment were from the two dominant ethnic groups in the region, Hazara and Tajik.

Our experimental design closely resembles the trust game originally proposed by Berg, Dickhaut, and McCabe (1995), in which economic gains can be realized when an investor exhibits trusting behavior by sending a positive amount of money to an anonymous trustee. The amount sent is tripled by the experimenter—representing gains to productivity—and the trustee receives the tripled amount and then chooses whether to return a portion of the profit to the investor. A self-interested investor will only send a positive amount if he expects the trustee to return an amount greater than her original investment. Since a similarly self-interested trustee will not return any money, the equilibrium for self-interested agents is that the investors send nothing. However, individual preferences, such as those related to altruism and reciprocity, often motivate trustees to return positive amounts, and consequently, investors send a positive amount (Camerer 2003). This game has been played in many settings with diverse subject pools, and the majority of investors do in fact send positive amounts and the majority of trustees return positive amounts.⁵

As Fehr and Rockenbach (2003), we implement two modifications to the basic trust game: first, investors request a desired back-transfer from the trustee. Second, the design includes a financial sanction that investors can choose to apply against trustees in case

⁴For example, Titmuss (1971) finds that voluntary blood donations decreased after a financial incentive was introduced. Similarly, financial incentives have been found to decrease individual willingness to accept a socially desirable, but locally unwanted nuclear waste storage facility (Frey and Oberholzer-Gee 1997). On the labor market, Bandiera, Barankay, and Rasul (2005) show that incorrectly designed pecuniary incentives can sometimes be detrimental to worker productivity when they negatively affect workers’ intrinsic motivations. In financial markets, Dupas and Robinson (2013) find that the introduction of a secure savings product in Kenya resulted in decreased participation in informal savings groups.

⁵See Johnson and Mislin (2011) for a comprehensive survey.

they fail to fulfill the desired back transfer. The sanction in our experiment is costless to the investor, but is costly to a trustee who fails to meet the investor’s request. We compare results in this game with the simple trust game, in which no sanction is available.

Absent of any state-dependent preferences (i.e. preferences “activated” by the presence of the sanction), applying sanctions should only improve mutual cooperation by disciplining trustees who fail to comply with the investor’s request. However, as demonstrated by previous research, economic incentives may crowd out pro-social motivations, such as altruism or reciprocity.⁶

Using a sample of German university students, Fehr and Rockenbach (2003) find that when investors apply sanctions, the amount returned by trustees decreases. Fehr and List (2004) use the same experimental design in Costa Rica and observe similar behavior both among a sample of students and another of CEOs. This phenomenon, which Bowles and Polania-Reyes (2012) term “categorical crowding out,” occurs when the behavioral effect of the sanction decreases trustworthiness, and the magnitude of this change is greater than the financial effect of the sanction (which should obviously increase back-transfers, *ceteris paribus*). When the option to impose a sanction is present but investors voluntarily abstain from using it, both Fehr and Rockenbach (2003) and Fehr and List (2004) find that this increases back-transfers relative to the trust game (i.e. when the sanction is not available). In this case, the investor sends a positive signal about her character by choosing to abstain from imposing the sanction, which activates trustees’ norms and preferences related to reciprocity. These studies imply that if cooperation or trustworthiness is already relatively high at the group level, the introduction and use of sanctions might actually lead to less efficient outcomes by decreasing the non-pecuniary motivation for cooperation.⁷

Our experiment allows us to examine how trust-based interactions are affected by the introduction of a specific type of institution: formal sanctions for failing to fulfill an implicit contract. In the experiment, investors decide how much of their endowment to

⁶See Bowles and Polania-Reyes (2012) for a review of literature on the effects of economic incentives on human intrinsic motivation.

⁷Other evidence shows that sufficiently weak sanctions can actually increase trustworthiness (Bohnet, Frey, and Huck 2001). Falk and Kosfeld (2006) show that a principal’s choice to control agents seems to negatively affect the agent’s effort compared to a situation in which the effort is not monitored, signaling the principal’s trust. Herold (2010) proposes a theoretical model in which principals are better-off by signing an incomplete contract with agents, as specifying too many contingencies in the contract might signal distrust, potentially leading to agents’ reduced performance. Psychological literature on young children shows that intrinsic motivations to perform a task diminish when an individual faces a controlling intervention (Deci and Ryan 1985).

transfer to trustees, what size back-transfer to request, and whether to impose a sanction on trustees for failing to comply with their request. Although trustees do not have the opportunity to negotiate or assent to the “offer” made by investors, the design can still be considered a contract of sorts, albeit one with an imbalance in negotiating power skewed towards the investor. Contracts between governments and citizens are similar in this regard.

Previous studies on contract enforcement and crowding out of moral incentives consider situations in which both agents belong to the same ethnic group, and in settings in which ethnic identity is not particularly salient. There is a great deal of evidence that other-regarding preferences towards members of one’s own social group—and ethnic group in particular—differ from those towards outsiders.⁸ We contribute to this field of research by demonstrating that ethnicity is also relevant to the ways in which individuals react to the introduction of a formal institution. Our results have practical implications for development policy in settings with informal local-level institutions that are limited to a specific ethnic group and in which formal institutions are relatively weak.⁹ We find that ethnicity does indeed matter, and that while sanctions do counter-act trustworthiness when individuals have a shared ethnic identity, the same is not true when individuals come from different ethnic groups. In fact, we find evidence that sanctions increase, or “crowd in”, behavioral trustworthiness in cross-ethnic interactions. This implies that the marginal effect of formal-contract enforcement mechanisms is higher when applied across ethnic groups, compared to within ethnic groups. The results are also important for the

⁸Bernhard, Fischbacher, and Fehr (2006) show that individuals in Papua New Guinea are more willing to expend resources on punishing norm violations when they are committed by someone from a different ethnic group towards their co-ethnics than when the victim is of a different group. In a study by Habyarimana et al. (2007), group network strength predicts the use of sanctions, which in turn fosters more cooperation among co-ethnic Ugandan slum dwellers. Alexander and Christia (2011) find that punishment as well as cooperation weaken among ethnoreligiously heterogeneous groups in post-war Bosnia-Herzegovina living in segregated communities. Interestingly, similarly heterogeneous groups are equally cooperative as homogenous groups when exogenously assigned to the integrated environment of a multi-ethnic high-school.

⁹For example, Fafchamps (2000) shows that access to informal credit in Zimbabwe is strictly limited to co-ethnic business partners, while formal credit institutions (i.e. banks) do not discriminate by ethnicity. Biggs, Raturi, and Srivastava (2002) examine access to credit in Kenya and find similar results. On a similar note, Lanjouw and Levy (2002) argue that informal land titling rules can, when well functioning, substitute for formal property rights. The authors thus propose that formal land titling programs should be mainly targeted at “young disorganized communities”; in other words in communities where informal rules and sanctioning mechanisms are lacking; ethnically heterogeneous societies being one such example. It should be noted that even if informal rules and sanctioning mechanisms exist, their existence might not be a first best from a welfare perspective: Field (2007) shows that granting formal land titles drives out existing social norms of community policing helping to prevent evictions in Peruvian settlements. Consequently, local populations were able to devote more hours to productive activities.

theoretical discussion of how group identity and social preferences are related.

In addition to introducing treatments based on ethnic identity, this study differs from previous experiments on state-dependent preferences in another important way: our subject pool has very little experience with formal, written, enforceable contracts. Overall, we find that financial incentives crowd out trustworthiness to a much lesser extent in our sample than in previous studies conducted on populations with higher levels of formal-market integration. Since institutions are rarely exogenous, an artefactual field experiment allows us to examine these issues with more precision and to make causal claims that would be much more challenging to study empirically using survey data. Our results help us to understand how social preferences may depend on the institutional context, and this has important policy implications for multi-ethnic settings with weak formal institutions.

In the next section, we lay out the design of the experiment, followed by a section outlining the theory behind our study; Section 4 presents results and Section 5 discusses the implications of our main findings.

2.2 Experimental Design

2.2.1 Experimental Games

To examine the effect of pecuniary sanctions on prosocial motivations we use two experimental games following the design of Fehr and Rockenbach (2003). There are two anonymously matched players in both games, an investor and a trustee, who both receive an initial endowment of $\omega = 100$ AFG, which was equivalent to around 2 USD at the time of the experiment. An investor, i , then chooses whether to “trust” the trustee by transferring some portion, $s_i \in [0, 10, 20, \dots, \omega_i]$ of his endowment. The amount sent is tripled by the experimenter, and the trustee receives $3s_i$. The trustee, t , then has the option of transferring some amount of what he receives, $r_t \in [0, 10, 20, \dots, 3s]$, back to the investor, thus sharing the benefits of the increased stake.

The payoffs for investors and trustees in the trust game, respectively, are:

$$\pi_i = \omega_i - s_i + r_t \tag{2.1}$$

$$\pi_t = \omega_t + 3s_i - r_t. \tag{2.2}$$

In contrast to a standard trust game, the investor also communicates a desired back-transfer, $r_i^* \in [0, 10, 20, \dots, 3s_i]$, to the trustee. In the baseline condition, this request is “cheap talk” and does not affect the payoffs of either party.

All subjects also played the sanctioning game, which adds one additional feature to the trust game with requested back-transfers: the investor can choose whether to impose a sanction, $f = 40$, dependent on whether the trustee’s back transfer is less than the amount requested by the investor. We denote the decision to impose the sanction as $p \in [0, 1]$, where $p = 1$ if the investor chooses to conditionally apply the sanction and zero otherwise.

The payoff of the function for the trustee in the sanctioning game is given by:

$$\pi_t = \omega_t + 3s_i - r_t - fp_t(\mathbf{1}\{r_t < r_i^*\}) \quad (2.3)$$

and the payoff for the investor is identical as in the trust game (Equation 2.1).

In the sanctioning game, with the parameters we use—which are identical to Fehr and Rockenbach (2003) and Fehr and List (2004)—the sanction is too small to allow the investor to capture the efficiency gains from a self-interested trustee in all but one, extreme case.¹⁰ However, assuming that the decision to impose the sanction does not negatively affect trustworthiness, there is no reason for an investor not to use it, as doing so provides a financial incentive in addition to pro-social motivations for trustees to transfer an amount at least as high as the investor’s request.

Investors also played a triple-dictator game, which resembles the trust game, but in which the trustee, in the role of a passive receiver in this game, has no option to return a portion of the amount received, and the investor, in the role of dictator, was aware of this when making his decision. As in the trust game, investors were given endowments equal to trustees, $\omega_t = \omega_i = 100$, and the amount transferred was tripled by the experimenter. The game allows us to identify altruistic motivations independently of the beliefs and strategic concerns that affect investors’ behavior in the trust and sanctioning games (Fershtman and Gneezy 2001; Cox 2004).

¹⁰If the investor sends $s_i = 10$, requests $r_i^* = 30$ and imposes the sanction, then the trustee will maximize his profit by returning $r_t = r_i^* = 30$ to avoid paying the fine $f = 40$. Thus, the maximum profit an investor can achieve when playing with a self-interested trustee is 10 AFG. Only one investor in our sample actually selected this strategy. If an investor sends $s_i = 20$ and requests $r_i^* = 40$, the trustee is indifferent between paying the fine and returning $r_t = r_i^* = 40$. If the trustee complies, the investor makes a profit of 20 AFG.

2.2.2 Treatments

In order to study how ethnicity affects trustworthiness and the response to sanctioning, we sampled subjects who identify as either Tajik or Hazara. These two groups live in mostly (unofficially) segregated communities, and while there has been no recent conflict directly between them, ethnicity is extremely salient in Afghan society. Moreover, most Tajiks are Sunni Muslims, while the majority of Hazaras are Shia, which means that they attend different mosques and celebrate different holidays. Since community leaders, or *kalantars*, serve small, ethnically homogeneous neighborhoods and are responsible for solving many local disputes, this also means that the majority of the informal and semi-formal institutions that affect day-to-day life for our subjects are also de facto segregated by ethnicity. Whenever a dispute spans across neighborhood boundaries, *the kalantars* representing the arguing parties are called in to resolve the issue. Given this, we expect these differences to be salient for group identity to play a role in decision-making.

Our approach differs from experiments using the "minimal group" paradigm originally applied in social psychology (Tajfel et al. 1971), and later widely adopted by economists (e.g., Charness, Rigotti, and Rustichini 2007; Heap and Zizzo 2009). This stream of literature exogenously establishes groups in a laboratory. The claim is that if differential treatment across groups occurs when ties are weak, then group behavior in general can be attributed to instinctive, rather than deeply rooted group-based norms. The artificial establishment of groups further allows researchers to make inference about causal effects of group membership. This is not possible with existing, endogenously established groups. However, Goette, Huffman, and Meier (2012) show that the behavior of members of established social groups and that of randomly established minimal groups differs substantially, especially in the context of enforcement of social norms.

We matched investors and trustees according to ethnic group such that the subject's partner in the game was either from the same ethnic group, the "*ingroup*" treatment, or from a different ethnic group, the "*outgroup*" treatment.

2.2.3 Procedure

In total we conducted 29 experimental sessions—14 with investors and 15 with trustees—with 442 subjects in October and November 2013 in 7 predominantly Tajik and 6 predominantly Hazara peri-urban areas of Mazar-e-Sharif, northern Afghanistan. The population is generally engaged in day labor or agriculture and communities are strongly ethnically

homogenous.

Subjects were randomly sampled according to their place of residence within the areas we selected. Individuals meeting our criteria (a married male 18 to 60 years of age, with at least one child, and of a particular ethnic group mobilized for a particular session) were invited to participate in the experiment. We studied males only, due to the cultural restrictions involved with working with female respondents in Afghanistan. The selection criteria were the same for the investors and the trustees and for ethnic groups. We were able to contact household heads in 76 percent of households visited, and 85 percent of those interviewed matched our criteria and were invited to participate. There is no significant difference in response rates across Tajik and Hazara communities (80 and 76 percent, respectively; $p=0.34$).

The experiment was conducted in groups of 15-20 subjects, who were informed that they would be matched with a person from a different community located in Mazar-e-Sharif, but that they would not know which community, specifically, nor would their partner be informed of their specific community. Subjects were read a short profile describing their partner, which included the general selection criteria used for subject sampling, in addition to the fact that their partner lived in a community that was “mostly Tajik” or “mostly Hazara” according to treatment.¹¹ This profile was read several times throughout the experiment and 90 percent of the subjects were able to recall the ethnicity of their partners at the end of the experimental session in an unprompted question asking about their partner characteristics. The treatment information was communicated during the group portion of the instructions, and thus our treatments are randomized at the session level. The other characteristics included in the profile remained constant for all treatments.

All subjects played both the trust game and the sanctioning game. We varied the order of the two games across sessions.¹² Following these two games, investors also played a triple dictator game and trustees were informed of the possibility that they would receive money from the investors’ dictator decisions.

Since the subject pool was largely illiterate, all instructions were given orally, using

¹¹The additional criteria were included to avoid an experimental demand effect that might result from making the aim of the study too obvious. Ethnicity is salient in everyday life in this region, and it is a reasonable assumption that including it in the description did not seem particularly out-of-place for subjects.

¹²In 75 percent of sessions trustees played the sanctioning game first, with the order reversed for the remaining sessions.

visual aids.¹³ After a general introduction of the experiment and explanation of the task in a group, the subjects were seated in private booths (See Figure 2.A.1) where they made their decisions using visual aids (See Figure 2.A.2) in privacy—though not anonymous to research assistants.¹⁴

We used a partial strategy method for trustees' decisions, in which subjects state contingent choices for several selected sets of combinations. Since it was impossible to obtain responses for the entire set of combinations of s , r^* , and p , we provided subjects with a set of four, exogenously given sets of parameters for the sanctioning game (including 2 with the sanction applied), and two sets of parameters for subjects for the trust game.¹⁵ Subjects were given the actual decisions made by their partners as well, though this was always the last decision made for each respective game. Subjects did not know ex-ante for which decision they would be paid.

There are three main advantages to employing this method. First, the parameters communicated to subjects were orthogonal to the group treatment. This allows us to study trustees' responses to each parameter directly. For example, if we were to examine trustees' responses to the use of the sanction without the strategy method, sanctioning would plausibly be correlated with both the group treatment as well as the amount sent and the amount requested, and would thus bias our estimates. Secondly, exogenously varying the parameters of the game gives us the potential to explore a range of possible decision types, even if those decisions were uncommonly chosen by investors. And third, the strategy method considerably increases statistical power.^{16,17}

¹³Our script builds on the instructions originally used in Barr (2003). See Online Appendix B for complete instructions.

¹⁴The two research assistants participating in all sessions were of Hazara and Tajik ethnic origin. Subjects within a session were matched with the research assistants randomly. The Tajik research assistant delivered the group instructions in all sessions.

¹⁵The parameters within each category were randomly selected from a pool of real decisions made by investors in earlier sessions or in practice rounds, which was the same for both treatments.

¹⁶We classify the allocation types based on two criteria: the level of investor's trust and the requested back-transfer. We consider an allocation as low-trust if the amount sent is 50 AFG or less and as high-trust otherwise. We label the decision *fair* in case the amount requested by the investor results in a distribution that is either equal for both subjects or advantageous for the investor (and *unfair otherwise*). Combining the two criteria allows us to assign each decision into one of four categories. In the trust game each trustee was presented with randomly selected strategy method decisions falling into two of these categories. In the sanctioning game, each trustee made four strategy method decisions in total, two decisions for each of the two categories he received in the trust game: one including the use of the sanction and one with no use of the sanction. Each decision within the same category had unique parameters. This procedure was used in order to limit within-subject variance.

¹⁷Despite the discussed advantages, decisions using the strategy method might differ from direct-response decisions observed in real life. Brandts and Charness (2011) survey the literature experimentally comparing the strategy and direct-response methods. They find that one out of five surveyed studies

At the end of every experimental session we administered a short, one-on-one survey to all subjects, including demographic information, membership in various formal and informal organizations, experience with formal and informal credit markets, and about their experience writing or signing formal contracts, and hypothetical questions designed to elicit their degree of experience with formal institutions. Key characteristics by treatment for investors and trustees are described in Table 2.1. The table also shows that our randomization into treatments was successful.¹⁸

Each subject received a 100 AFN show-up fee. This is a substantial amount of money, compared to wages for a day of manual labor of around 150 AFN. Subjects were informed that the payoff they earned in the games would be distributed in two days to allow us to match their responses with their partners’.

2.3 Theoretical Background

In this section we present a theoretical framework for interpreting the effects of the sanction on trustees’ decisions. Our goal in this study is to determine whether sanctions crowd out (or crowd in) trustworthy behavior, and if so, whether such an effect differs between the *ingroup* and *outgroup* treatments. The sanction, however, can be expected to influence the amount returned by trustees in two distinct ways: by changing the trustee’s payoff function (i.e. the financial effect of the sanction), and by activating certain state-dependent preferences that affect trustworthiness. To our knowledge, the most complete model that best illustrates possible unintended effects of the availability and of the use of formal sanctions is the model of state-dependent preferences in Bowles and Polania-Reyes (2012). Rather than making our own theoretical contribution, we present a slightly modified version of the theoretical framework of Bowles and Polania-Reyes (2012) for the tractability of presented results.

using a trust game find a difference and one finds mixed evidence. Three remaining studies find no differences in both amounts sent and—reassuringly for this paper—in amounts returned. The paper finding a difference reports substantially higher amounts returned using the direct-response method, so the results presented here might represent a lower bound. Unfortunately, none of the studies surveyed in Brandts and Charness (2011) examine the exact type of game we use here.

¹⁸A strongly significant treatment difference arises only in investors’ responses to a World Values Survey (WVS) question asking "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?" Reassuringly, the main investor results hold when splitting the sample by the response to this question (see Appendix Table 2.A.5). In all regressions we control for an index combining responses to the three last questions reported in Table 2.1, among other characteristics. Appendix Table 2.A.1 breaks down the summary statistics by ethnicity and shows that the difference in responses to the WVS questions is mainly driven by the Tajik participants.

Based on Bowles and Polania-Reyes (2012), we assume that the trustee's utility in the trust and sanctioning games is influenced by a combination of material and other-regarding preferences:

$$U_t(\pi_t, \alpha_{ti}\pi_i) \quad (2.4)$$

where π_t represents trustee t 's payoff function depending on the game. Likewise, π_i is investor i 's payoff from Equation 2.1. The term α_{ti} represents the trustee's other regarding preferences towards the investor, such that:

$$\alpha_{ti} = \overbrace{\beta_t^g}^{\text{altruism}} + \overbrace{p_i\bar{\lambda}_t^g + n_i\underline{\lambda}_t^g}^{\text{behavioral effect of sanction}}, \quad (2.5)$$

where β_t^g captures t 's preferences for i 's payoff, conditional on whether t and i have a shared group affiliation $g \in \{ingroup, outgroup\}$, but not conditional on whether the sanction was available or used.¹⁹ When $\beta_t^g > 0$, this term captures t 's altruism towards i .

We include two parameters to capture state-dependent preferences: the parameter $\bar{\lambda}_t^g$ represents a set of the trustee's state-dependent preferences that change with application of the sanction, $p_i \in \{0, 1\}$. Again, we allow this parameter to vary by group treatment, g . The parameter $\bar{\lambda}_t^g$ encompasses several motivations. Firstly, the presence of the sanction might fundamentally change the nature of the relationship between the trustee and the investor, and thus activate a different set of preferences. Secondly, since the investor chooses whether or not to apply the sanction, it may signal something about his character or intentions, and this in turn may change the weight given to his payoff in the trustee's utility.²⁰ When $\bar{\lambda}_t^g < 0$, the sanction crowds out trustworthiness. If the degree of crowding out is large enough, then the sanction could cause a trustee to return less than he would without the sanction (categorical crowding out). It is also possible that $\bar{\lambda}_t^g > 0$, which would indicate that the financial incentive of the sanction crowds in trustworthiness. This would be the case if the sanction reinforces existing norms or social preferences.

We also include the term $n_i\underline{\lambda}_t^g$, where $n_i \in \{0, 1\}$ is equal to one if the sanction was available, but the investor chose not to impose it (in the trust game, $n = p = 0$). The

¹⁹This model assumes that the amount sent by the investor and the requested back transfer are held constant. It is likely that these parameters meaningfully interact with sanctioning as well. However, we omit them here for simplicity. Moreover, the strategy method we employ, in which sanctioning is orthogonal to the other parameters, allows us to consider the effect of sanctioning independently.

²⁰See Houser et al. (2008), who study the difference between these two motivations.

parameter λ_t^g captures both the change in preferences that results from the contextual difference in the relationship between the sanctioning and the trust games, as well as the implicit signal of trust that is communicated by voluntary abstention from sanctioning. This “good news” about the investor’s beliefs and intentions likely increases the amount returned by trustees (Bowles and Polania-Reyes 2012).

While we would ideally wish to estimate the values of both λ parameters for the *ingroup* and *outgroup* treatments, this is not possible given that the financial and behavioral effects of the sanction go hand in hand. However, our results do allow us to make inferences about the direction of the effects, and their relative values across treatments, which we examine in the next section.

2.4 Results

We begin by analyzing how both the availability and use of sanctions affect the amount returned in the trust game across the *ingroup* and *outgroup* treatments. This allows us to test our main hypothesis: group identity plays a role in how individuals react to sanctioning, and that consequently the efficiency of formal institutions might differ between ethnically homogeneous and heterogeneous settings. In the following sub-section, we go deeper into the behavioral motivations underlying trustees’ reactions to the sanctions between treatments, and consider how the fairness of requests might influence response to the sanction across treatments in subsection 2.4.3, before discussing investors’ behavior in subsection 2.4.4.

2.4.1 Trustee Experimental Results

We begin by analyzing the amount returned by trustees in each game and treatment. In addition to comparing averages across treatments, we analyze behavior in the games using the following econometric model:

$$\begin{aligned}
ShareReturned_{ji} = & \beta_0 + \beta_1 Ingroup_i + \beta_2 Sanctioning_{ji} + \beta_3 Ingroup_i \times Sanctioning_{ji} \\
& + \beta_4 NoSanctioning_{ji} + \beta_5 Ingroup_i \times NoSanctioning_{ji} \\
& + \mathbf{E}_{ji}'\boldsymbol{\Gamma} + \mathbf{X}_i'\boldsymbol{\Delta} + \varepsilon_{ji}
\end{aligned} \tag{2.6}$$

where $ShareReturned_{ji}$ is the percentage of the total amount received that trustee i returns for decision j , (i.e. $r_t/3s_i$), $Ingroup_i$ is a treatment dummy, $Sanctioning_{ij}$ and $NoSanctioning_{ij}$ are indicators for the application of the sanction, with the trust game as the omitted category; \mathbf{E}_{ij} is a set of experimental controls, including the *amount sent* and *share requested* by the investor; \mathbf{X}_{ij} is a set of individual characteristics,²¹ and ε_{ij} is the error term, clustered at the individual level.

Table 2.2 and Figure 2.1 summarize the results for trustees in the trust and sanctioning games. We limit our analysis to the decisions made by trustees using the strategy method, in which the amount sent, the requested back-transfer, and the decision to apply the sanction were exogenously assigned to trustees by the experimenter, and the parameters were on average the same, independent of the group treatment. This allows us to draw causal inferences about the effect of group treatments in relation to each individual parameter.

The first two columns of Figure 2.1 demonstrate that trustees return a significantly higher portion of what they receive in the *ingroup* treatment than in the *outgroup* treatment. When trustees were paired with an investor from the same ethnic group, the *share returned* is, on average, 58.16 percent of the total amount they received from investors, compared to only 43.26 percent in the *outgroup* treatment, which is statistically significant ($p < 0.01$).²² This indicates that ethnicity is indeed salient among the population that we study, and has an effect on trustworthiness. In column 1 of Table 2.3, we report results from Equation 2.6. The *ingroup* treatment is associated with an increase of 15.41 percentage points in the *share returned* in the trust game ($p < 0.01$).

Next, we compare trustees' behavior in the sanctioning game. Columns 3 and 4 of Figure 2.1 report the shares returned by subjects in the *ingroup* and *outgroup* treatments, respectively, in the *sanctioning* condition. Compared to the trust game, subjects in the *outgroup* treatment send back much more in the *sanctioning* condition, returning 58.74 percent on average—an increase of 15.48 percentage points (see Table 2.2). The difference is highly significant ($p < 0.01$). In contrast, for the *ingroup* treatment there is only a slight increase in the amount returned relative to the trust game (3.25 percentage points), and the difference is not significant ($p = 0.22$). Notably, under the *sanctioning* condition, the

²¹The set of individual characteristics includes the trustee's ethnicity, age, number of household members, a dummy for literacy, years spent living continuously in Mazar-e-Sharif, log of individual monthly income (in AFN), a dummy for whether the individual had ever signed a contract and an index of perceptions of trust and fairness towards others (combining answers to three World Values Survey questions).

²²All significance levels reported for comparison of means are from simple two-sided mean comparison t-tests.

gap between *ingroup* and *outgroup* treatments narrows to only 2.67 percentage points, which is no longer significant ($p=0.32$).

Columns 5-6 of Figure 2.1 present the results from the *no sanctioning* condition. For trustees in the *outgroup* treatment, the *share returned* in the *no sanctioning* condition falls between the levels for the *trust* and *sanctioning* conditions at 49.83 percent, a decrease of 8.90 percentage points over the *sanctioning condition* ($p<0.01$) and 6.58 percentage points higher than in the trust game ($p<0.01$).

For the *ingroup* treatment, however, trustees returned an average of 57.87 percent in the *no sanctioning* condition, which is virtually identical to the share returned in the trust game ($p=0.91$), and only 3.54 percentage points lower than the share returned in the *sanctioning* condition, which is also not statistically significant ($p=0.19$).

The regression results are similar. In column 1 of Table 2.3, we compare results from the *sanctioning* and *no sanctioning* conditions with trust game results, and add interactions with the group treatment. Since we include an interaction term for *ingroup*sanctioning*, the coefficient for *sanctioning* represents the effect of sanctions for the *outgroup* treatment. For those in the *outgroup* treatment, the *sanctioning* condition increases the *share returned* by 13.96 percentage points over the corresponding trust game result ($p<0.01$). The *ingroup*sanctioning* interaction is significant ($p<0.01$) and of virtually the same magnitude, -13.44, indicating no effect of the *sanctioning* condition over the trust game for the *ingroup* treatment. The effect of the *no sanctioning* condition is similar, though smaller in magnitude for the *outgroup* treatment, with a coefficient of 5.29 ($p=0.02$). Again, the *ingroup*no sanctioning* interaction is of virtually the same magnitude, -5.47 ($p=0.05$) indicating no effect of the *no sanctioning* condition in the *ingroup* treatment.

In columns 2-3 of Table 2.3, we split the sample by treatment and compare the *sanctioning* and *no sanctioning* conditions with trust game results. Neither the *sanctioning* ($p=0.99$), nor the *no sanctioning* ($p=0.82$) conditions differ significantly from the trust game for the *ingroup* treatment (column 2). The differences for the *outgroup* treatment, however, are larger and statistically significant (column 3). Relative to the trust game, the *share returned* increases under the *sanctioning* condition by 13.98 percentage points ($p<0.01$) and by 5.48 percentage points under the *no sanctioning* condition ($p=0.01$).²³

²³The results are strongest when using a subsample of sessions when the sanctioning game was played before the trust game. See Appendix Tables 2.A.2 and 2.A.3. Since the behavior in the sanctioning game was of greatest interest for us, we oversampled this game ordering by 3 to 1 compared to the situation when the trust game was played first.

Our results differ from those of previous studies using similar designs (Fehr and Rockenbach 2003; Fehr and List 2004) in that the *sanctioning* condition does not lower the average amount returned. While this does not necessarily indicate that either the *sanctioning* or *no sanctioning* conditions do not crowd out moral incentives, it does suggest that any such effect is substantially smaller than is present in previous studies.

Additionally, we find that the sanction has a higher marginal effect in cross-ethnic situations than in ethnically homogenous ones. This is true when comparing the *sanctioning* condition with both the trust game and *no sanctioning* conditions.

The fact that we do not see a response to the sanction for subjects in the *ingroup* treatment does not necessarily indicate that there are no underlying effects, as the financial effect of the sanction may cancel out the behavioral effect. In fact, there are several scenarios consistent with our results: First, the *sanctioning condition* could influence behavior by crowding out moral incentives, but less so in the *outgroup* treatment. Second, the *sanction* could complement or “crowd in” non-financial incentives, but more so for the *outgroup*. And third, the *sanction* could have an opposite behavioral effect in each group treatment, reinforcing pro-social norms for the *outgroup*, but crowding out moral incentives for the *ingroup*. Understanding this underlying mechanism would make it possible to draw broader conclusions from the results of the experiment, and we attempt to distinguish between these possibilities in the following sub-section.

2.4.2 The Behavioral Effect of Sanctions

While we are not able to obtain direct estimates of the parameters in Equation 2.5, we are able to make assumptions about the underlying mechanisms that drive the main results by breaking down trustees’ decisions into three categories, according to whether the amount returned was greater than, equal to, or less than the amount requested by the investor. As before, we consider only decisions made using the strategy method to avoid endogeneity among the parameters.

Absent any state-dependent preferences, one would expect the frequency of subjects who return more than the requested amount in the *sanctioning* condition to be the same as in the trust game. To see this, refer to Equation 2.5. If $\bar{\lambda}_t^g = 0$, and if a trustee’s utility is maximized by returning $r > r^*$ in the trust game, then—holding other parameters constant—the introduction of the sanction provides no financial incentive to change one’s behavior, since neither the trustee’s nor investor’s payoff would be affected

given the status quo. On the other hand, if sanctions crowd out moral incentives, (i.e. if $\bar{\lambda}_t^g < 0$), then the trustee will maximize his utility by returning a smaller amount when the sanction is present, since he now puts less weight on the investor's payoff. If the magnitude of $\bar{\lambda}_t^g$ were large enough, holding all other parameters constant, the trustee would no longer return more than the requested amount, and therefore on average we would expect to see a drop in the frequency of trustees returning $r > r^*$ in the *sanctioning* condition, relative to the trust game. On the other hand, if the sanction reinforces, or crowds in, existing norms of reciprocity or altruism (i.e. if $\bar{\lambda}_t^g > 0$), then the frequency of decisions in which $r > r^*$ could be larger in the *sanctioning* condition than in the trust game, following similar logic.

The frequencies of each type of decision in the trust game and *sanctioning* and *no sanctioning* conditions, by group treatment, are shown in Figure 2.2. In the trust game, 31.21 and 16.76 percent of subjects returned more than the requested amount in the *ingroup* and *outgroup* treatments, respectively. In the *ingroup* treatment, the frequency of decisions in which the amount returned was more than the requested amount drops by 8.27 percentage points in the *sanctioning* condition relative to the trust game. This suggests that sanctions do in fact crowd out trustworthiness in the *ingroup* treatment. The difference is marginally significant ($p=0.09$). In the *outgroup* treatment, however, there is actually an increase of 4.71 percentage points in the frequency of decisions in which the amount returned exceeds the amount requested in the *sanctioning* condition relative to the trust game, though the difference is not statistically significant ($p=0.26$). There is virtually no change in the number of subjects returning more than requested between the trust game and *no sanctioning* condition in either the *ingroup* and *outgroup* treatments ($p=0.68$ and $p=0.73$, respectively).²⁴

Next, we turn to the case in which an individual returns less than the requested amount in the trust game, $r < r^*$. Holding other variables constant, a trustee's behavior might change with the introduction of the sanction in response to both the financial effects and the behavioral effects, which makes predictions less straightforward. Whether the individual will return less than the requested amount depends on the relative size of $(\beta_t^g + \bar{\lambda}_t^g)$. With a smaller value of $\bar{\lambda}_t^g$, one would expect to see a greater frequency of

²⁴The results of a probit model presented in columns 1 to 3 in Table 2.A.4 directionally match the findings presented in the main text (controlling for observables). While the increase in the frequency of returning more than requested in the *sanctioning* condition, relative to the trust game, is statistically significant for the *outgroup*, the decrease in similar decisions for the *ingroup* is not statically significant. Also, after controlling for observables, we find an increase in the frequency of returning more than requested in the *no sanctioning* condition, relative to the trust game, for the outgroup treatment only.

decisions in which less was returned than the requested amount. With a larger value of β_t^g , the sanction should be less effective, since trustees will be more willing to sacrifice their own payoff to avoid the socially harmful sanction.

In Figure 2.2, we observe that there is a precipitous drop in the frequency of decisions in which trustees return less than the requested amount from the trust game to the *sanctioning* condition in the *outgroup* treatment by nearly half, from 56 percent to 29 percent, ($p < 0.01$). In the *ingroup* treatment, there is also a decrease in the frequency of those returning less than requested, yet the decrease is comparatively smaller: from 41 in the trust game to 31 percent in the *sanctioning* condition ($p = 0.06$).²⁵

In the *outgroup* treatment in the *no sanctioning* condition trustees returned less than the requested amount in 59 percent of decisions, which is roughly the same as in the trust game, and the difference between the *sanctioning* and *no sanctioning* conditions is statistically significant ($p < 0.01$). In the *ingroup* treatment, trustees returned less than the requested amount in 35.4 percent of decisions in the *no sanctioning* condition, which is not statistically different from the *sanctioning* condition, ($p = 0.41$).

To summarize, these results suggest financial sanctions crowd out trustworthiness for those in the *ingroup* treatment. For the *outgroup*, however, it appears that sanctions actually reinforce behavioral motivations to return a higher amount. For the *ingroup*, the sanction has little to no effect on the frequency of decisions in which trustees return less than the requested amount. This implies that $(\beta_t^g + \bar{\lambda}_t^g)$ is lower for the *ingroup* treatment than the *outgroup* treatment, which would be consistent with either of the two following scenarios: i) less altruism or preferences for social efficiency in the *ingroup* treatment or ii) that sanctions influence behavior differently for the *ingroup* and *outgroup* treatments, such that $\bar{\lambda}_t^{ingroup} > \bar{\lambda}_t^{outgroup}$. Since previous literature (e.g., Bernhard, Fischbacher, and Fehr 2006) suggests that there should be more altruism in the *ingroup* treatment, the latter explanation seems more likely.

2.4.3 Sanctioning and Fairness

Both theory (Rabin 1993; Fehr and Schmidt 1999) and experimental evidence (e.g., Herrmann, Thöni, and Gächter 2008; Henrich et al. 2010) suggest that individuals tend to reward fair behavior and punish behavior that is considered unfair. Fehr and Rockenbach (2003) find that responses to the *sanctioning* condition differ between “fair” and “unfair”

²⁵The results are confirmed using a probit estimation that includes observables. See columns 4 to 6 of Table 2.A.4.

requests. Here, we also use their definition of a fair request: the amount sent and requested by the investor is such that the payoffs of the investor and the trustee are either equal or are in favor of the trustee, which implies that $r^*/3s \leq 0.67$.²⁶

In columns 1-3 of Table 2.4 we present results for fair requests only (as before, considering only strategy-method decisions). Coefficients in column 1 are directionally similar to the results using all decisions in column 1 of Table 2.3. However, the net effect of the *sanctioning* condition on *ingroup* trustees is now negative. This can be seen more clearly in column 2, which includes only fair requests for *ingroup* subjects: the effect of the *sanctioning* condition, relative to the trust game, reduces the share returned by 5.24 percentage points ($p=0.02$), indicating that sanctions crowd out trustworthiness, and that the negative behavioral effect of the sanction is larger in magnitude than the financial effect (categorical crowding out). We also find stronger results for the *no sanctioning* condition for *ingroup* subjects when we consider only fair requests: there is a corresponding decrease of 5.29 percentage points in the share returned relative to the trust game ($p=0.01$). Interestingly, there is no significant difference between the *sanctioning* and *no sanctioning* conditions ($p=0.98$).

For the *outgroup*, the effect is reversed, and the *sanctioning* condition has a similar effect on the sub-sample of fair decisions as it does overall, increasing the share returned by 9.64 percentage points ($p<0.01$). Likewise, and contrary to the results for the *ingroup* treatment, the *no sanctioning* condition is associated with a 6.23 percentage-point increase in the share returned ($p=0.01$), relative to the trust game, for fair decisions in the *outgroup* treatment (Table 2.4, column 3). As with the *ingroup* treatment, there is no statistically significant difference between *sanctioning* and *no sanctioning* for *outgroup* subjects when requests are fair ($p=0.20$).

Next, we turn to unfair allocations, in which the amount requested leaves the investor with a higher payoff than the trustee. For unfair requests, the *sanctioning* condition increases back-transfers for both the *ingroup* and *outgroup* treatments, though in column 4 of Table 2.4, we find that the effect is much smaller in the *ingroup* treatment ($p=0.02$), and in column 5, we find that the effect of *sanctioning* relative to the trust game, while positive (5.43 percentage points), is not statistically significant for the *ingroup* treatment ($p=0.20$); nor is there any statistically significant difference between the *sanctioning* and

²⁶Note that we assumed in our design, ex ante, that trustee's decisions might qualitatively differ with respect to this dimension, and we provided subjects an equal number of each type of strategy-method decisions, giving us a roughly balanced number of observations in each category. In total we have 547 observations for fair requests and 452 unfair requests.

the *no sanctioning* conditions ($p=0.96$).

For the *outgroup* treatment, there is a larger effect of *sanctioning* on the share returned when requests are unfair, with back-transfers increasing by 18.56 percentage, relative to the trust game ($p<0.01$). There is no statistically significant effect of the no sanctioning condition relative to the trust game ($p=0.28$).

These results indicate that the reaction to the *sanctioning* and *no sanctioning* conditions does, indeed, differ according to the fairness of requests. In the *ingroup* treatment, when requests are fair, we find clearer evidence that both the *sanctioning* and *no sanctioning* conditions crowd out trustworthiness, indicating that both $\bar{\lambda}^{ingroup} < 0$ and $\underline{\lambda}^{ingroup} < 0$. As discussed in the following section, the vast majority of decisions made by investors in both treatments, for both the trust game and the sanctioning game, can be classified as fair. It is therefore possible that trustees had less clear interpretations of the investor's intentions when faced with unfair decisions, and indeed this is reflected by the larger standard errors for most coefficients when trustees faced unfair requests in columns 4-6 compared to fair requests in columns 1-3 of Table 2.4. For the *outgroup* treatment, our main findings are consistent regardless of whether the request was fair or unfair.

2.4.4 Investor Experimental Results

We now turn to investors' decisions in the trust, sanctioning and triple-dictator games, which are presented in Table 2.5.²⁷ In the trust game, investors sent 57.21 AFN and 56.19 AFN on average in the *ingroup* and *outgroup* treatments, respectively ($p=0.76$). The lack of a statistically significant difference in amounts sent is puzzling, as investors might plausibly anticipate higher levels of trustworthiness in the *ingroup* treatment and would thus achieve Pareto dominant outcomes by sending more.

However, investors in the *ingroup* treatment request higher back-transfers, 51.97 versus 42.46 percent of the tripled amount sent $r_i^*/3s_i$, in the *ingroup* and *outgroup* treatments, respectively ($p<0.01$). This is consistent with investors attempting to increase profits without decreasing their minimum guaranteed payoff, as would be the case if an investor sent more. Investors' beliefs—revealed using an incentivised question—support this interpretation, as expected returns are higher in the *ingroup* treatment compared to the *outgroup* (48 vs. 42 percent of the amount received by trustees, respectively,

²⁷A graphical representation of the investor results is also presented in the Appendix Figure 2.A.3. In this figure we split the results by treatment, game, and whether the sanction was applied or not.

$p=0.04$).²⁸ Because efficiency gains in the trust game are proportional to money sent, this means that *ingroup* interactions are not more efficient than *outgroup* interactions, despite the higher levels of trustworthiness. As discussed in the previous subsection, the majority of decisions, 88 percent, are classified as fair. The higher share of the transfer requested by investors in the *ingroup* treatment leads to a lower percentage of fair requests in the *ingroup* treatment compared to the *outgroup* (84 vs. 93 percent, $p=0.06$).

In the sanctioning game, despite the higher marginal return to sanctioning that we see among trustees in the *outgroup* treatment, we find no treatment difference in the use of the sanction: 36.54 percent of investors in the *ingroup* treatment and 38.10 percent of investors in the *outgroup* treatment chose to apply the sanction—the difference is not statistically significant ($p=0.83$). Overall, the majority of those who applied the sanction in both treatments expected trustees to comply with their request. In the *outgroup* treatment, 39.47 percent of those who applied the sanction expected trustees to be fined; compared to 28.13 percent of those who used the sanction in the *ingroup*, the difference is not statistically significant, $p=0.33$.

We do not observe a statistically significant difference in amounts sent across treatments in the sanctioning game either. Investors sent on average 55.96 AFN in the *ingroup* treatment and 56.67 AFN in the *outgroup* treatment ($p=0.82$). As in the trust game though, requested back transfers are higher in the *ingroup* treatment: 54.41 percent, compared to 48.97 percent of the total amount received by trustees in the *outgroup* treatment. The difference is marginally significant ($p=0.08$). Neither is the difference between the amounts sent in the trust and sanctioning games significant for either treatment: $p=0.60$ and $p=0.82$ for the *ingroup* and *outgroup* treatments, respectively. This indicates that the introduction of the sanctioning mechanism did not lead to increased efficiency for either group treatment. Requested back transfers were higher, however, in the sanctioning game (see Table 2.5). Subjects in the *outgroup* treatment requested 6.51 percentage points more of the tripled amount sent in the sanctioning game than in the trust game ($p=0.02$). The difference is smaller for the *ingroup* treatment, at 2.69 percentage points, and is not statistically different from zero ($p=0.30$),

Although the difference between the *ingroup* and *outgroup* treatments in requested shares is significant ($p=0.08$), the magnitude is smaller compared to the difference ob-

²⁸Generally, investors expected their requests in the trust game to be met or exceeded: 66.34 and 66.67 percent in *ingroup* the *outgroup* treatments, respectively. The treatment difference is not statistically significant ($p=0.96$).

served in the trust game. Investors' beliefs again reflect the lower expected trustworthiness of *outgroup* trustees compared to the *ingroup* (51 vs. 45 percent, $p < 0.01$).

Surprisingly, in the triple dictator game we also fail to find a statistically significant difference in allocations: subjects sent on average 44.23 AFN in the *ingroup* treatment, and sent 46.73 AFN in the *outgroup* treatment ($p = 0.58$). As one would predict, dictator allocations are lower than allocations in both the trust and sanctioning games for treatments at the 99 percent level.²⁹

Interpreting parameters in the games independently is problematic, since investors simultaneously chose an amount to send, request back and, in the sanctioning game, whether to impose the sanction. Because of this, it is difficult to draw conclusions from each component of the decisions separately, as the investor's intentions and expectations in, say, requesting a large back transfer might be quite different if he also made a large transfer, compared to requesting an equivalent portion back from a smaller transfer. Another way of analyzing the results, which mitigates this problem, is to consider the overall payoff that an investor's allocation and requested back transfer implies for himself ($\omega_i - s + r^*$) and for his partner ($\omega_t + 3s - r^*$). We can then compare this across games and treatments to obtain a measure of how requested profit changed exogenously, according to the setting. In the trust game, the requested profit of the *ingroup*, 130.67 AFN, is significantly higher than in the *outgroup* treatment, 112.86 AFN ($p < 0.01$). While the requested profit might represent different strategies according to the game, investors also reported that they believed they would earn more in the trust game in the *ingroup* treatment, 122.12 AFN, than in the *outgroup* treatment, 111.31 AFN ($p = 0.06$). *Ingroup* subjects also expected their partners to earn less, 103.47 AFN, compared to 111.94 AFN in *outgroup* treatments ($p = 0.13$).

The same pattern holds for the sanctioning game: the requested profit in the *ingroup* treatment was 134.81 AFN and 124.17 in the *outgroup* treatment. The difference is marginally significant ($p = 0.10$). The same holds for expected profit, which was 126.80 and 117.50 AFN in the *ingroup* and *outgroup* treatments, respectively ($p = 0.07$).

In Appendix Table 2.A.6, we regress expected and requested profit for the investor and his partner, for all three games, including dummies for treatment and the sanctioning and dictator game (the excluded category is the trust game). We also include the

²⁹The differences in differences between dictator allocations and trust game allocations by treatment are not statistically significant. The same is true when comparing dictator and sanctioning game allocations (available upon request).

same set of individual controls included in Equation 2.6. This allows an intuitive and simple method of examining differences between treatments across all three games, while avoiding the endogeneity problems that arise when analyzing parameters of decisions in the experiment individually. The results generally confirm our analysis using comparison of means. *Ingroup* subjects requested and expected higher profits for themselves and requested and expected their partners to earn less in both the trust and sanctioning games. There is no statistically significant interaction term between treatment and game, except for the difference in requested profits between the trust and dictator games. Those in the *ingroup* treatment, for whom requested profits in the trust and dictator game were further apart, reflecting higher requested back-transfers in the trust game for investors in the *ingroup* treatment.

2.5 Discussion

The results that we have presented indicate that the behavioral effect of economic sanctions on trustworthiness does indeed differ according to ethnic identity: pecuniary sanctions are not as effective in increasing back-transfers in the *ingroup* treatment as they are in the *outgroup* treatment. Sanctions crowd out behavioral trustworthiness in the *ingroup* treatment. Our results suggest that the adverse effect of sanctions on behavioral trustworthiness is present to a much lesser extent when individuals are from different ethnic groups, and that sanctions may in fact reinforce norms for cooperation, or crowds in, behavioral trustworthiness in inter-ethnic interactions (i.e. $\bar{\lambda}_t^{ingroup} < 0 < \bar{\lambda}_t^{outgroup}$). When we consider only fair requests—those that leave the trustee at least as well off as the investor—the results are stronger, and we find evidence of categorical crowding out for the *ingroup* treatment: the decrease in trustworthiness in response to the sanction is larger in magnitude than the financial effect of the sanction.

In the *no sanctioning* condition, we see an increase in back transfers, relative to the trust game, but only for the *outgroup* treatment. For the *ingroup*, we do not see any effect of the *no sanctioning* condition, relative to the trust game. In terms of the model presented in Section 3, this indicates that $\lambda_t^{outgroup} > 0$. Recall that this parameter captures state-dependent preferences connected both to the situation—sanctions are possible—as well as the intentions communicated when an investor intentionally refrains from using the sanction when it is available to him. One plausible explanation for this trend is that those in the *ingroup* treatment react negatively to the situation—the mere presence of

sanctions, which defines subjects' roles in the game differently than in the trust game—and this cancels out any positive intentions communicated by an investor who refrains from using the sanction when it is available. Moreover, when we consider only fair requests, we find a decrease in back-transfers in the *no sanctioning* condition, relative to the trust game, which further supports this finding. In fact, we find no difference between the *sanctioning* and *no sanctioning* conditions for the *ingroup* treatment when we consider fair requests. This strongly suggests that it is the context of the sanctioning game that crowds out behavioral trustworthiness in the *ingroup* treatment, rather than the intentions communicated by the actions of the investor.

These differences in state-dependent preferences towards *ingroup* and *outgroup* members could result from a few underlying mechanisms. Firstly, it is possible that we observe more crowding out of behavioral trustworthiness in the *ingroup* treatment simply because there is more trustworthiness to crowd out. Potentially there is a diminishing marginal return on the behavioral effect of sanctions, and this is smaller in the *outgroup* treatment since trustworthiness is lower in general. However, this seems unlikely given our results, due to the magnitude of the effect. In the *sanctioning* condition, we see that the gap between *ingroup* and *outgroup* subjects virtually disappears, and in fact we find evidence that sanctions actually increase trustworthiness. If the treatment difference in response to sanctions were driven by initial levels of trustworthiness alone—rather than a difference in state-dependent preferences—this would not be the case.

Another possibility is that there are different perceptions of fairness between the *ingroup* and *outgroup* treatments, and this drives the differences in responses to the *sanctioning* and *no sanctioning* conditions. Potentially, state-dependent preferences are independent of group identity, but do depend on whether the request is considered fair. If a given request is considered fair by the trustee when the investor is from a different ethnic group, but unfair when the investor is a co-ethnic, then we would expect a different reaction to the sanctioning conditions, even if the underlying preferences related to sanctioning were identical. To this end, we test several definitions of fair requests and do not observe any clear pattern in response to different thresholds in either treatment. Coefficients from regressions on sub-samples defined by the requested back transfer are presented in Appendix Figure 2.A.4. If differences in perceptions of fairness alone were driving the treatment difference, we would expect to see a similar effect of the *sanctioning* and *no sanctioning* conditions on the *ingroup* and *outgroup* treatments after adjusting the definition of fairness. The fact that we see no such pattern indicates that this is not

the case.

Finally, we are left with the interpretation that group identity affects behavioral responses to sanctions in a more fundamental way. Our results suggest that state-dependent, other-regarding preferences differ according to group identity. This complements previous findings that group identity, and ethnic identity in particular, plays an important role in defining other-regarding preferences (Fershtman and Gneezy 2001; Bernhard, Fischbacher, and Fehr 2006).

In general, the subjects in our experiment exhibit behavior that differs from previous studies using similar games in that we observe that sanctions crowd out trustworthiness to a lesser degree than in previous experiments (Fehr and Rockenbach 2003; Fehr and List 2004). Our sample subjects have very little experience interacting with formal institutions and this may underlie the difference in behavior, suggesting that experience and culture may play a role in shaping preferences relating to institutional settings. The fact that our sample behaves differently than those in previous studies might also call into question the generalizability of our findings. However, we do find similar behavior in the two ethnic groups that we study, and our results are not limited to either Hazara or Tajik subjects (see Appendix Tables 2.A.7 and 2.A.8). While this is reassuring, we acknowledge that these two ethnic groups share many cultural similarities and live in the same setting. It is possible that the relationship between formal sanctioning mechanisms and ethnic identity would differ in other settings, and more research should be done in this area, to understand these cultural differences.

While we do not find that investors apply the sanction more frequently in the *outgroup* treatment, as might be expected given the trustee results, there are several possible explanations. First of all, investors' behavior is not only strategic, and it is possible that preferences exist for applying the sanction, and that these are sensitive to group identity. For example, if subjects conjectured that sanctions would be more effective in increasing back transfers in the *outgroup* treatment, but also cared more about punishing (perceived) norm violators from their own group—even if they believed it would be costly for them, by decreasing back transfers in expectation—then this could potentially lead to equal frequencies of *sanctioning* in each treatment. On the other hand, perhaps some investors in the *outgroup* treatment refrain from sanctioning due to fear of damaging relations with the other ethnic group.

Additionally, we work with a population that has little experience with formal institutions: only around 6 percent of the total sample have ever signed a contract. It is possible

that this lack of experience with sanctioning means that investors' beliefs were not accurate, and given the one-shot nature of the experiment, they were unable to predict trustee behavior precisely enough in each treatment to respond accordingly. Examining how such behavior changes with increased exposure to formal institutions is an intriguing area for future study.

Understanding how the particular preferences that we consider in this article differ by group identity is of particular interest because it relates to the role of ethnicity in the adoption of formal institutions. Informal institutions play a much greater role in developing economies, are usually organized at the community level, and are ethnically homogeneous. There is evidence from the laboratory (Habyarimana et al. 2007) and field (Miguel and Gugerty 2005) that increased ethnic diversity makes the provision of public goods more difficult. Our results indicate that the introduction of a particular type of formal institution has a higher marginal return when applied to cross-ethnic interactions compared to interactions between co-ethnics. This suggests that efforts to build new formal institutions might be advised to concentrate on inter-ethnic settings. Our findings might also help to explain patterns in the development of formal institutions throughout history as a mechanism for enforcing cooperation in multi-ethnic societies.

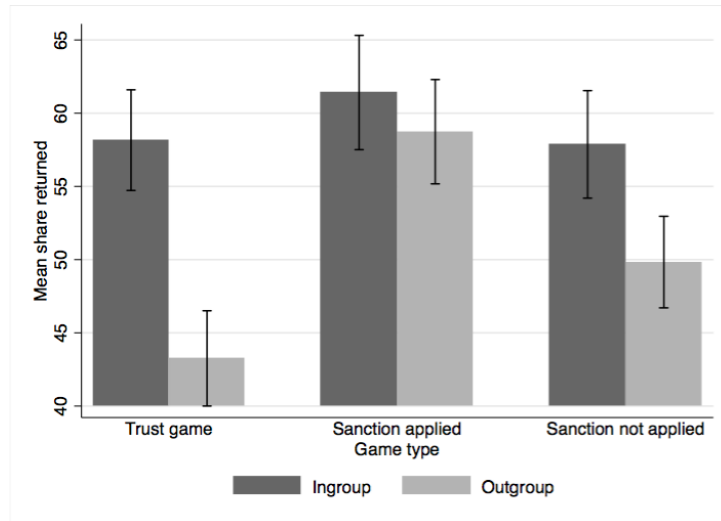


Figure 2.1: Trustees' Average Share Returned by Game and Treatment.

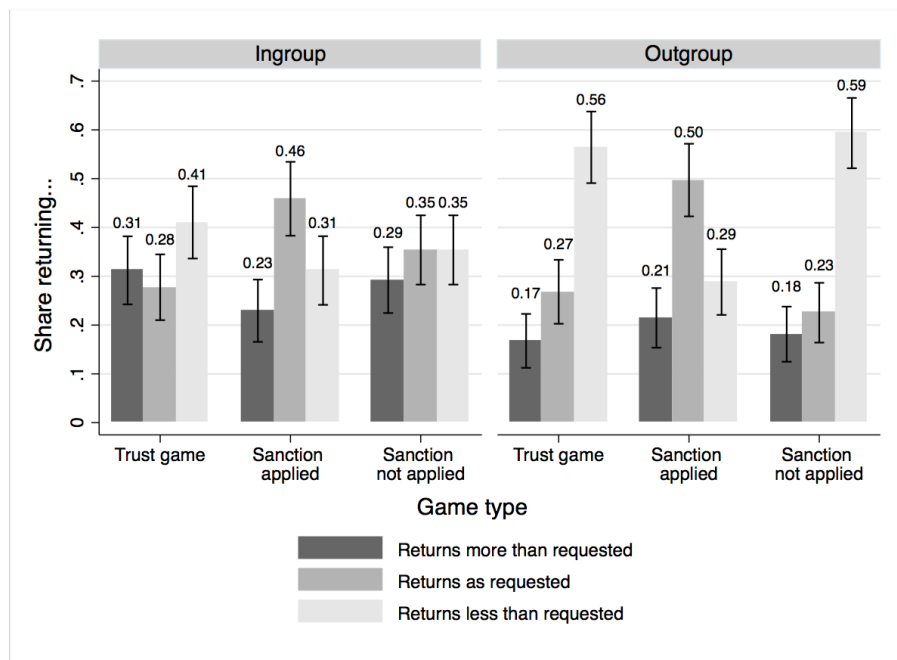


Figure 2.2: Trustees' Decisions to Meet Investors' Requests by Game and Treatment.

Table 2.1: Individual Characteristics by Role and Treatment

<i>Sample</i>	<i>Investors</i>			<i>Trustees</i>		
	<i>Ingroup</i> (1)	<i>Outgroup</i> (2)	Difference (1)-(2) (T-test p-value) (3)	<i>Ingroup</i> (4)	<i>Outgroup</i> (5)	Difference (4)-(5) (T-test p-value) (6)
Age	41.04 (14.00)	41.11 (13.23)	-0.07 (0.97)	38.94 (12.98)	36.87 (12.95)	2.08 (0.29)
Household members	7.74 (3.16)	7.49 (2.91)	0.25 (0.56)	8.17 (3.13)	8.55 (5.61)	-0.37 (0.59)
Can read a letter (d)	0.32 (0.47)	0.29 (0.46)	0.03 (0.61)	0.40 (0.49)	0.51 (0.50)	-0.11 (0.14)
Years living in Mazar	13.81 (13.53)	11.67 (12.93)	2.14 (0.26)	19.21 (15.04)	16.34 (15.93)	2.87 (0.22)
∞ Individual's monthly income (AFN)	1609.27 (3844.60)	1627.53 (3139.47)	-18.25 (0.97)	1773.61 (3633.12)	24119.10 (211848.10)	-22345.50 (0.33)
Ever written contract in the past (d)	0.08 (0.27)	0.06 (0.23)	0.02 (0.53)	0.06 (0.24)	0.03 (0.18)	0.03 (0.42)
Others can be trusted (d)	0.54 (0.50)	0.75 (0.44)	-0.21*** (0.00)	0.52 (0.50)	0.61 (0.49)	-0.08 (0.27)
Others are fair (d)	0.39 (0.49)	0.38 (0.49)	0.02 (0.80)	0.28 (0.45)	0.33 (0.47)	-0.05 (0.47)
Others are selfish (d)	0.70 (0.46)	0.58 (0.50)	0.12* (0.08)	0.70 (0.46)	0.69 (0.47)	0.01 (0.86)
Observations	112	88		86	89	

Note: Means reported in Columns 1, 2, 4, and 5. Standard deviations in parentheses. Columns 3 and 6 report the difference in means between the *ingroup* and the *outgroup* treatment. P-values of a two-sided t-test are reported in parentheses in columns 3 and 6. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. (d) denotes a dummy variable. The high income for trustees in the *outgroup* treatment is driven by one individual who reported an income of 2000000 AFN. The last three rows show responses to World Values Survey questions asking: 1) "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?", 2) "Do you think most people would try to take advantage of you if they got a chance, or would they try to be fair?", and 3) "Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves?"

Table 2.2: Trustees' Behavior in Games: Aggregate and by Treatment, Strategy Method Allocations

<i>Sample</i>	<i>Total</i> (1)	<i>Ingroup</i> (2)	<i>Outgroup</i> (3)	Difference (2)-(3) (T-test p-value) (4)
Trust game				
Share returned	50.58 (23.65)	58.16 (22.90)	43.26 (22.05)	14.90*** (0.00)
Observations	352	173	179	
Sanctioning game				
<i>Sanctioning condition</i>				
Share returned	60.05 (24.89)	61.41 (25.77)	58.74 (24.00)	2.67 (0.32)
Observations	347	170	177	
<i>No sanctioning condition</i>				
Share returned	53.81 (23.45)	57.87 (24.81)	49.83 (21.37)	8.04*** (0.00)
Observations	360	178	182	

Note: Means reported in Columns 1-3. Standard deviations in parentheses. Strategy method allocations only. Column 4 reports the difference in means between the *ingroup* and the *outgroup* treatment. P-values of a two-sided t-test are reported in parentheses in column 4. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level.

Table 2.3: Effect of Sanctions on Share Returned in Trust and Sanctioning Games Across Treatments

<i>Sample</i>	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>
Dependent variable	Share returned		
	(1)	(2)	(3)
Ingroup	15.41*** (2.71)		
Sanctioning condition	13.96*** (2.47)	-0.01 (2.15)	13.98*** (2.50)
Ingroup x Sanctioning condition	-13.44*** (3.22)		
No sanctioning condition	5.29** (2.17)	-0.39 (1.78)	5.48** (2.17)
Ingroup x No sanctioning condition	-5.47* (2.78)		
Amount sent (AFN)	-0.15*** (0.03)	-0.14*** (0.04)	-0.18*** (0.04)
Share requested	0.37*** (0.04)	0.48*** (0.06)	0.28*** (0.04)
Control variables	Yes	Yes	Yes
Constant	23.19*** (5.91)	31.41*** (10.88)	31.98*** (7.09)
Observations	999	491	508
R-squared	0.27	0.29	0.26
F-test			
H_0 : Sanctioning equals no sanctioning			
<i>Ingroup</i> p-value	0.68	0.83	
<i>Outgroup</i> p-value	0.00		0.00

Note: OLS coefficients. Standard errors in parentheses (clustering at individual level). *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. Strategy method allocations only. In each regression we control for the trustee's ethnicity, age, number of household members, a dummy for literacy, years spent living continuously in Mazar-e-Sharif, log of income (AFN), a dummy for whether the individual has ever signed a contract and an index of perceptions of trust and fairness towards others (3 questions). The F-test compares the sanctioning and no sanctioning condition coefficients.

Table 2.4: Effect of Sanctions on Share Returned Across Treatments, by Fairness (OLS)

Sample	<i>Fair requests</i>			<i>Unfair requests</i>		
	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>
	Share returned					
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)
Ingroup	13.98*** (3.09)			18.02*** (4.53)		
Sanctioning condition	9.08*** (2.87)	-5.24** (2.24)	9.64*** (2.93)	19.53*** (3.95)	5.43 (4.23)	18.56*** (4.08)
Ingroup x Sanctioning condition	-13.59*** (3.63)			-14.06** (5.78)		
No sanctioning condition	5.61** (2.41)	-5.29** (2.07)	6.23** (2.38)	5.23 (4.07)	5.28 (3.64)	4.41 (4.04)
Ingroup x No sanctioning condition	-10.68*** (3.19)			-0.33 (5.30)		
Amount sent (AFN)	-0.12*** (0.04)	-0.18*** (0.05)	-0.10* (0.05)	-0.18*** (0.04)	-0.13** (0.06)	-0.27*** (0.06)
Share requested	0.39*** (0.05)	0.50*** (0.06)	0.31*** (0.07)	0.27*** (0.10)	0.43*** (0.15)	0.16 (0.12)
Constant	25.21*** (6.41)	31.69** (13.65)	33.90*** (8.18)	27.57** (12.02)	36.57** (16.89)	37.36** (15.46)
Observations	547	265	282	452	226	226
R-squared	0.23	0.27	0.22	0.20	0.14	0.23
F-test						
H_0 : Sanctioning equals no sanctioning						
<i>Ingroup</i> p-value	0.72	0.97		0.77	0.98	
<i>Outgroup</i> p-value	0.35		0.00	0.19		0.20

Note: OLS coefficients. Standard errors in parentheses (clustering at individual level). *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. Strategy method allocations only. In each regression we control for the trustee's ethnicity, age, number of household members, a dummy for literacy, years spent living continuously in Mazar-e-Sharif, log of income (AFN), a dummy for whether the individual has ever signed a contract and an index of perceptions of trust and fairness towards others (3 questions). The F-test compares the sanctioning and no sanctioning condition coefficients.

Table 2.5: Investors' Behavior in Games: Aggregate and by Treatment

<i>Sample</i>	<i>Total</i> (1)	<i>Ingroup</i> (2)	<i>Outgroup</i> (3)	Difference (1)-(2) (T-test p-value) (4)
Trust game				
Amount sent	56.76 (22.46)	57.21 (23.67)	56.19 (20.99)	1.02 (0.76)
Share requested back	0.48 (0.22)	0.52 (0.22)	0.42 (0.20)	0.1*** (0.00)
Belief about share returned	0.45 (0.21)	0.48 (0.22)	0.42 (0.20)	0.06** (0.04)
Share of fair allocations	0.88 (0.33)	0.84 (0.37)	0.93 (0.26)	-0.09* (0.06)
Requested profit in AFN (investor)	122.71 (42.09)	130.67 (42.86)	112.86 (39.17)	17.82 (0.00)
Requested profit in AFN (trustee)	190.80 (57.09)	183.75 (54.33)	199.52 (59.51)	-15.77 (0.06)
Expected profit in AFN (investor)	117.29 (38.90)	122.12 (39.13)	111.31 (37.98)	10.81 (0.06)
Expected profit in AFN (trustee)	196.22 (57.58)	192.31 (56.98)	201.07 (58.29)	-8.76 (0.30)
Sanctioning game				
Amount sent	56.28 (20.91)	55.96 (21.97)	56.67 (19.66)	-0.71 (0.82)
Share requested back	0.52 (0.21)	0.54 (0.22)	0.49 (0.20)	0.05* (0.08)
Belief about share returned	0.48 (0.19)	0.51 (0.20)	0.45 (0.17)	0.06** (0.02)
Sanction applied	0.37 (0.48)	0.37 (0.48)	0.38 (0.49)	-0.02 (0.83)
Share of fair allocations	0.85 (0.36)	0.79 (0.41)	0.92 (0.28)	-0.13** (0.02)
Requested profit in AFN (investor)	130.05 (43.71)	134.81 (48.35)	124.17 (36.61)	10.64 (0.10)
Requested profit in AFN (trustee)	182.50 (53.41)	177.12 (53.38)	189.17 (53.01)	-12.05 (0.12)
Expected profit in AFN (investor)	122.62 (34.95)	126.80 (37.16)	117.50 (31.50)	9.30 (0.07)
Expected profit in AFN (trustee)	190.00 (53.71)	185.24 (54.59)	195.83 (52.35)	-10.59 (0.18)
Dictator game				
Amount sent	45.16 (25.57)	44.23 (26.50)	46.31 (24.48)	-2.08 (0.58)
Observations	188	104	84	

Note: Means reported in Columns 1, 2, and 3. Standard deviations in parentheses. Column 4 reports the difference in means between the *ingroup* and the *outgroup* treatment. P-values of a two-sided t-test are reported in parentheses in column 4. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level.

2.A Appendix 2



Figure 2.A.1: Participants in an Individual Session in a Pop-up Field Laboratory.



Figure 2.A.2: Decision-Making Environment With Visual Aids Used for Experimental Sessions.

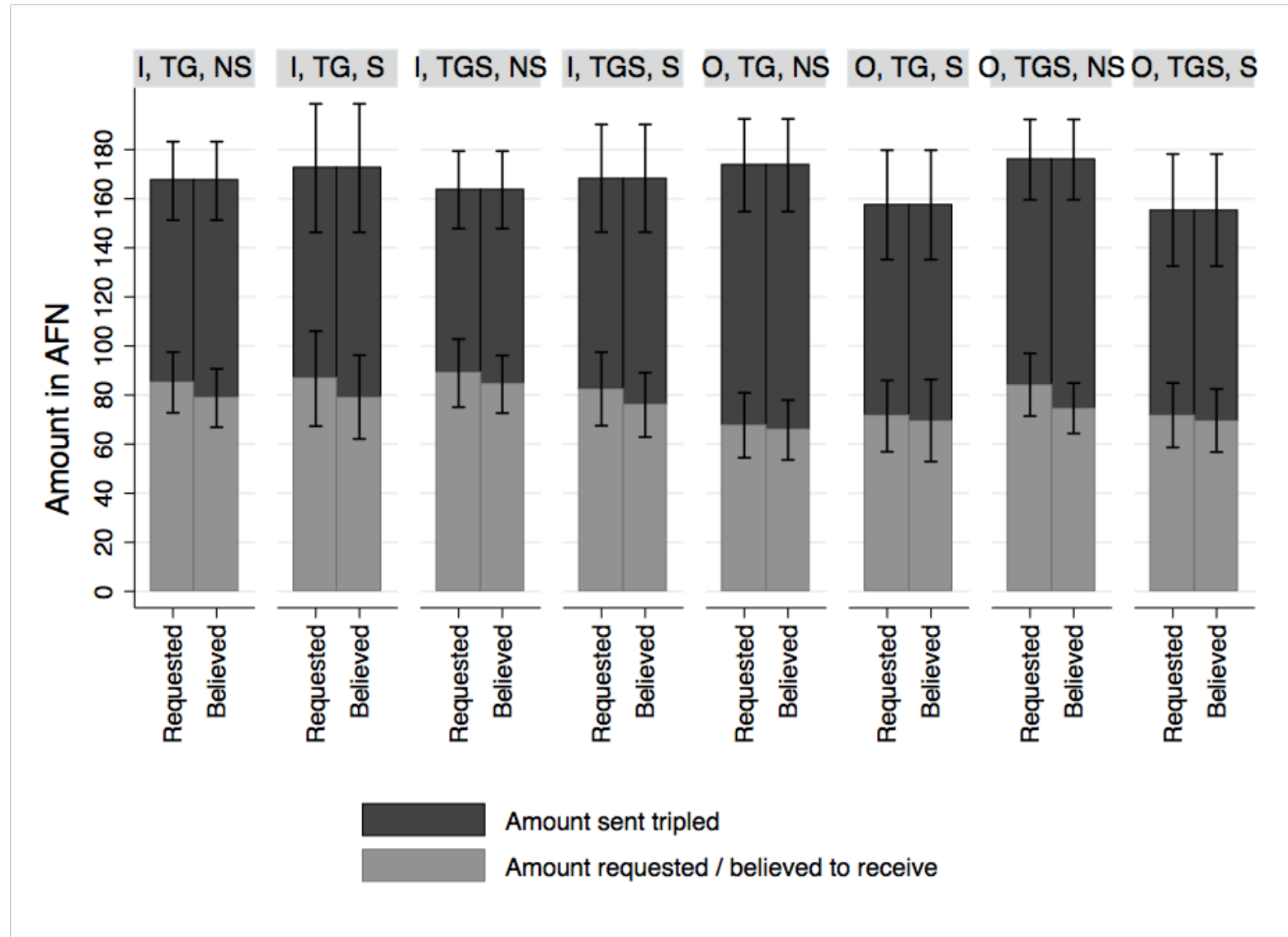


Figure 2.A.3: Investors' Expected Amount Returned Relative to Requests by Choice of Sanction

Note: The dark grey bars represent amounts sent in AFN. The light grey color represents amounts of AFN requested back (left in each pair) and amount believed to receive back (right in each pair). The abbreviations used in descriptions of bars are as follows:

I = Ingroup, O = Outgroup, TG = Trust game, TGS = Sanctioning game, NS = No sanctioning, S = Sanctioning. The lines represent 95 percent confidence intervals.

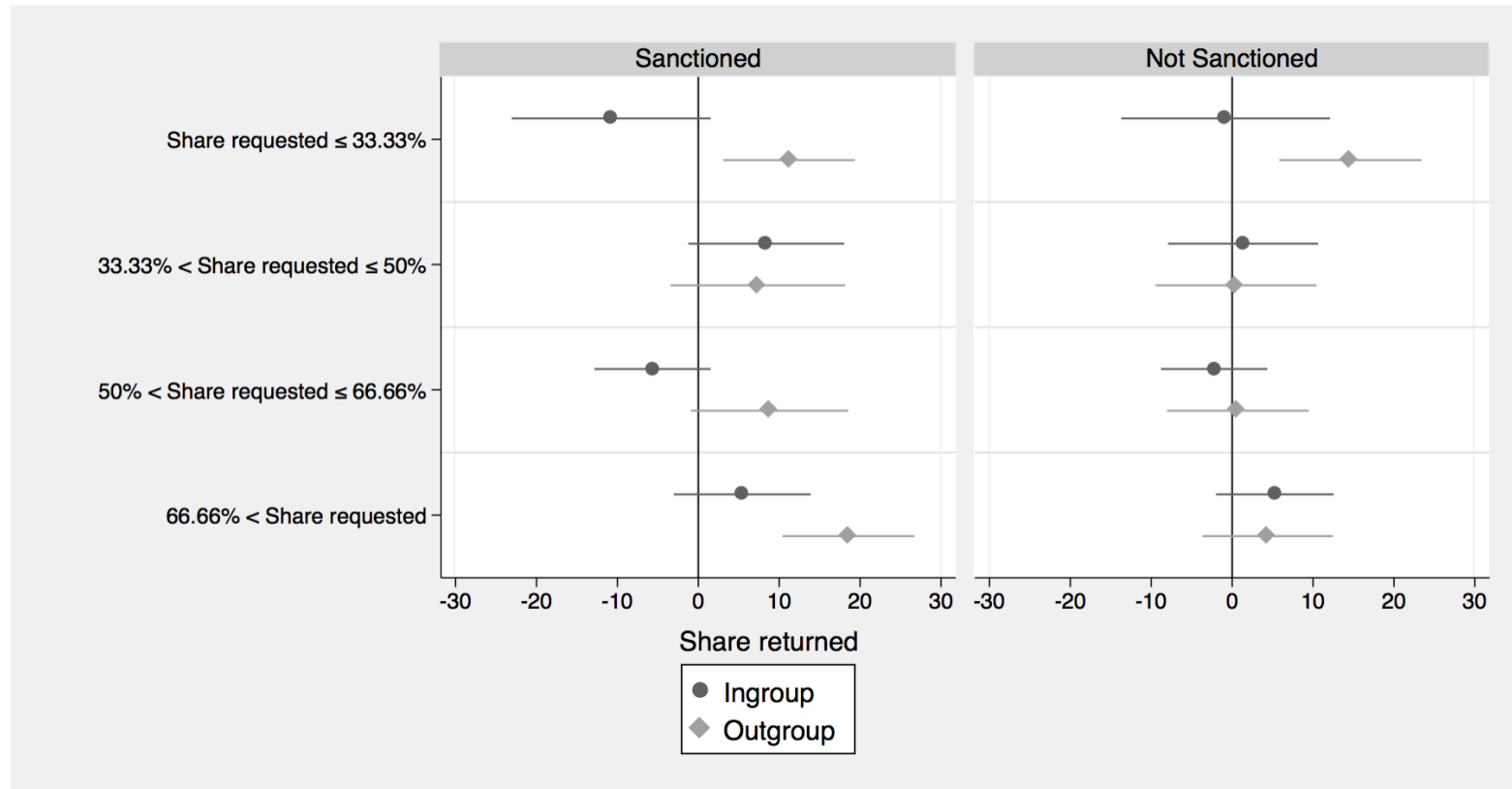


Figure 2.A.4: Coefficient Plot: Share Returned Regressed on *Sanctioning* and *No Sanctioning* Conditions, by Different Ranges of Share Requested.

Note: Regression results from Equation 2.6. The dependent variable is the share returned in the sanctioning or no sanctioning conditions. Each mark represents the coefficient for the given treatment, for a range of requests (implying various thresholds for what might be considered fair requests). The lines represent 95 percent confidence intervals.

Table 2.A.1: Individual Characteristics by Role, Ethnicity and Treatment

<i>Sample</i>	<i>Investors</i>					
	<i>Tajik</i>			<i>Hazara</i>		
	<i>Ingroup</i>	<i>Outgroup</i>	Difference (1)-(2) (T-test p-value)	<i>Ingroup</i>	<i>Outgroup</i>	Difference (4)-(5) (T-test p-value)
	(1)	(2)	(3)	(4)	(5)	(6)
Age	40.46 (15.44)	41.28 (12.97)	-0.82 (0.80)	41.52 (12.84)	41.00 (13.54)	0.52 (0.84)
Household members	8.20 (3.33)	8.22 (2.97)	-0.02 (0.97)	7.37 (2.99)	6.98 (2.79)	0.39 (0.48)
Can read a letter (d)	0.42 (0.50)	0.34 (0.48)	0.08 (0.48)	0.24 (0.43)	0.25 (0.44)	-0.01 (0.92)
Years living in Mazar	23.10 (13.90)	17.92 (14.72)	5.19 (0.10)	6.47 (7.36)	7.35 (9.47)	-0.88 (0.58)
Individual's monthly income (AFN)	2601.85 (5304.23)	2147.22 (3721.02)	454.63 (0.66)	843.57 (1836.29)	1274.53 (2654.51)	-430.96 (0.29)
Ever written contract in the past (d)	0.18 (0.39)	0.09 (0.28)	0.09 (0.22)	0.00 (0.00)	0.04 (0.19)	-0.04 (0.12)
Others can be trusted (d)	0.44 (0.50)	0.81 (0.40)	-0.37*** (0.00)	0.63 (0.49)	0.71 (0.46)	-0.08 (0.36)
Others are fair (d)	0.42 (0.50)	0.19 (0.40)	0.23*** (0.03)	0.37 (0.49)	0.50 (0.50)	-0.13 (0.17)
Others are selfish (d)	0.74 (0.44)	0.47 (0.51)	0.27*** (0.01)	0.67 (0.47)	0.66 (0.48)	0.01 (0.90)
Observations	50	36		62	52	

Continued on next page

Table 2.A.1 – continued from previous page

<i>Sample</i>	<i>Trustees</i>					
	<i>Tajik</i>			<i>Hazara</i>		
	<i>Ingroup</i> (1)	<i>Outgroup</i> (2)	Difference (1)-(2) (T-test p-value) (3)	<i>Ingroup</i> (4)	<i>Outgroup</i> (5)	Difference (4)-(5) (T-test p-value) (6)
Age	38.88 (13.13)	35.38 (15.59)	3.51 (0.27)	39.00 (12.98)	38.08 (10.34)	0.92 (0.71)
Household members	9.07 (3.52)	10.10 (7.71)	-1.03 (0.44)	7.32 (2.45)	7.31 (2.53)	0.01 (0.98)
Can read a letter (d)	0.33 (0.48)	0.58 (0.50)	-0.24** (0.03)	0.45 (0.50)	0.45 (0.50)	0.01 (0.96)
Years living in Mazar	29.29 (13.47)	25.23 (18.59)	4.06 (0.26)	9.59 (8.98)	9.08 (8.04)	0.51 (0.77)
Individual's monthly income (AFN)	1806.67 (2807.85)	1256.25 (3054.57)	550.42 (0.40)	1742.05 (4309.60)	42782.65 (285448.00)	-41040.61 (0.34)
Ever written contract in the past (d)	0.08 (0.28)	0.03 (0.16)	0.06 (0.27)	0.05 (0.21)	0.04 (0.21)	0.00 (0.96)
Others can be trusted (d)	0.48 (0.51)	0.68 (0.47)	-0.20* (0.07)	0.57 (0.50)	0.55 (0.50)	0.02 (0.87)
Others are fair (d)	0.26 (0.45)	0.41 (0.50)	-0.15 (0.16)	0.30 (0.46)	0.27 (0.45)	0.03 (0.75)
Others are selfish (d)	0.69 (0.47)	0.68 (0.47)	0.02 (0.88)	0.70 (0.46)	0.69 (0.47)	0.01 (0.91)
Observations	42	40		44	49	

Note: Means reported in Columns 1, 2, 4, and 5. Standard deviations in parentheses. Columns 3 and 6 report the difference in means between the *ingroup* and the *outgroup* treatment. P-values of a two-sided t-test are reported in parentheses in columns 3 and 6. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. (d) denotes a dummy variable. The high income for Hazara trustees in the *outgroup* treatment is driven by one individual who reported an income of 2000000 AFN. The last three rows show responses to World Values Survey questions asking: 1) "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?", 2) "Do you think most people would try to take advantage of you if they got a chance, or would they try to be fair?", and 3) "Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves?"

Table 2.A.2: Effect of Sanctions on Share Returned Across Treatments by Order of Games

<i>Sample</i>	<i>Sanctioning Game Played First</i>			<i>Sanctioning Game Played Second</i>		
	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>
	Share returned					
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)
Ingroup	18.64*** (2.94)			5.32 (4.90)		
Sanctioning condition	16.93*** (2.91)	-1.93 (2.31)	17.17*** (2.98)	4.20 (3.52)	3.87 (4.37)	3.05 (4.12)
Ingroup x Sanctioning condition	-17.66*** (3.73)			-0.36 (5.48)		
No sanctioning condition	5.96** (2.50)	-1.93 (1.98)	6.31** (2.52)	3.09 (4.40)	3.19 (3.40)	2.07 (4.55)
Ingroup x No sanctioning condition	-7.39** (3.18)			0.02 (5.55)		
Amount sent (AFN)	-0.09*** (0.03)	-0.11** (0.05)	-0.10** (0.04)	-0.28*** (0.06)	-0.24** (0.09)	-0.42*** (0.10)
Share requested	0.42*** (0.04)	0.60*** (0.05)	0.30*** (0.05)	0.20*** (0.07)	0.23** (0.10)	0.25** (0.10)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Constant	17.23*** (5.72)	27.69*** (8.83)	21.06*** (7.71)	48.73* (27.34)	44.49 (44.13)	58.32*** (14.73)
Observations	753	359	394	246	132	114
R-squared	0.35	0.44	0.29	0.25	0.30	0.41

Continued on next page

Table 2.A.2 – continued from previous page

<i>Sample</i>	<i>Sanctioning Game Played First</i>			<i>Sanctioning Game Played Second</i>		
	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>
Dependent variable	Share returned					
	(1)	(2)	(3)	(4)	(5)	(6)
	F-test					
	H ₀ : Sanctioning equals no sanctioning					
<i>Ingroup</i> p-value	0.73	1.00		0.82	0.84	
<i>Outgroup</i> p-value	0.00		0.00	0.82		0.85

Note: OLS coefficients. Standard errors in parentheses (clustering at individual level). *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. Strategy method allocations only. In each regression we control for the trustee's ethnicity, age, number of household members, a dummy for literacy, years spent living continuously in Mazar-e-Sharif, log of income (AFN), a dummy for whether the individual has ever signed a contract and an index of perceptions of trust and fairness towards others (3 questions). The F-test compares the sanctioning and no sanctioning condition coefficients.

Table 2.A.3: Effect of Sanctions on Share Returned Across Treatments, by Fairness and Order of Games (OLS)

<i>Sample</i>	<i>Sanctioning game first</i>					
	<i>Fair requests</i>			<i>Unfair requests</i>		
	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>
	Share returned					
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)
Ingroup	15.08*** (3.33)			23.74*** (4.92)		
Sanctioning condition	12.07*** (3.33)	-6.62*** (2.34)	12.51*** (3.52)	22.11*** (4.65)	3.40 (4.56)	21.74*** (4.75)
Ingroup x Sanctioning condition	-18.04*** (4.10)			-18.32*** (6.60)		
No sanctioning condition	7.24** (2.78)	-6.79*** (2.45)	7.79*** (2.75)	4.07 (4.14)	3.92 (3.75)	3.97 (4.22)
Ingroup x No sanctioning condition	-13.98*** (3.77)			-0.46 (5.44)		
Amount sent (AFN)	-0.10** (0.04)	-0.17*** (0.05)	-0.06 (0.06)	-0.11** (0.04)	-0.06 (0.06)	-0.16*** (0.06)
Share requested	0.43*** (0.05)	0.60*** (0.07)	0.33*** (0.07)	0.31*** (0.11)	0.47** (0.18)	0.22* (0.13)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Constant	24.25*** (5.95)	32.16*** (11.89)	25.49*** (8.52)	16.35 (12.30)	35.84** (16.89)	19.06 (16.18)
Observations	414	195	219	339	164	175
R-squared	0.29	0.42	0.25	0.28	0.20	0.26

Continued on next page

Table 2.A.3 – continued from previous page

<i>Sample</i>	<i>Sanctioning game second</i>					
	<i>Fair requests</i>			<i>Unfair requests</i>		
	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>
	Share returned					
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)
	F-test					
	H_0 : Sanctioning equals no sanctioning					
<i>Ingroup</i> p-value	0.75	0.95		0.95	0.88	
<i>Outgroup</i> p-value	0.12		0.13	0.00		0.00
<i>Continued on next page</i>						

Table 2.A.3 – continued from previous page

<i>Sample</i>	<i>Sanctioning game second</i>					
	<i>Fair requests</i>			<i>Unfair requests</i>		
	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>
	Share returned					
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)
Ingroup	6.39 (6.69)			2.37 (9.05)		
Sanctioning condition	-3.47 (6.15)	0.01 (3.84)	-1.27 (6.18)	8.95 (7.39)	7.00 (9.69)	10.01 (8.93)
Ingroup x Sanctioning condition	4.06 (7.68)			-1.73 (11.49)		
No sanctioning condition	-1.41 (5.63)	-0.11 (3.20)	1.23 (5.08)	4.89 (11.53)	6.30 (8.42)	5.34 (14.21)
Ingroup x No sanctioning condition	1.14 (6.47)			1.62 (13.63)		
Amount sent (AFN)	-0.18** (0.08)	-0.25** (0.11)	-0.21* (0.11)	-0.34*** (0.08)	-0.19** (0.09)	-0.53*** (0.12)
Share requested	0.24** (0.09)	0.35** (0.16)	0.23 (0.14)	0.21 (0.22)	0.23 (0.29)	0.38 (0.35)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Constant	53.01 (39.31)	56.61 (60.15)	70.46*** (24.11)	29.95 (35.00)	1.86 (43.62)	19.20 (46.32)
Observations	133	70	63	113	62	51
R-squared	0.23	0.41	0.34	0.32	0.24	0.52

Continued on next page

Table 2.A.3 – continued from previous page

<i>Sample</i>	<i>Sanctioning game second</i>					
	<i>Fair requests</i>			<i>Unfair requests</i>		
	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>
Dependent variable	Share returned					
	(1)	(2)	(3)	(4)	(5)	(6)
	F-test					
	H_0 : Sanctioning equals no sanctioning					
<i>Ingroup</i> p-value	0.83	0.98		0.90	0.90	
<i>Outgroup</i> p-value	0.71		0.67	0.60		0.60

Note: OLS coefficients. Standard errors in parentheses (clustering at individual level). *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. Strategy method allocations only. In each regression we control for the trustee's ethnicity, age, number of household members, a dummy for literacy, years spent living continuously in Mazar-e-Sharif, log of income (AFN), a dummy for whether the individual has ever signed a contract and an index of perceptions of trust and fairness towards others (3 questions). The F-test compares the sanctioning and no sanctioning condition coefficients.

Table 2.A.4: Effect of Punishment on Amount Returned Relative to Request (Probit)

<i>Sample</i>	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>
Dependent variable	Returned more than requested			Returned less than requested		
	(1)	(2)	(3)	(4)	(5)	(6)
Ingroup	0.21*** (0.05)			-0.18*** (0.06)		
Sanctioning condition	0.15*** (0.06)	-0.05 (0.05)	0.13** (0.05)	-0.33*** (0.05)	-0.13** (0.05)	-0.36*** (0.06)
Ingroup x Sanctioning condition	-0.14*** (0.04)			0.20** (0.09)		
No sanctioning condition	0.09* (0.05)	0.00 (0.05)	0.08** (0.04)	0.02 (0.06)	-0.09* (0.05)	0.01 (0.06)
Ingroup x No sanctioning condition	-0.07 (0.05)			-0.11 (0.07)		
Amount sent (AFN)	-0.00 (0.00)	-0.00* (0.00)	-0.00 (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)
Share requested	-0.01*** (0.00)	-0.01*** (0.00)	-0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	999	491	508	999	491	508
F-test						
H_0 : Sanctioning equals no sanctioning						
<i>Ingroup</i> p-value	0.29	0.26		0.34	0.43	
<i>Outgroup</i> p-value	0.20		0.22	0.00		0.00

Note: Marginal effects reported for probit regressions. Standard errors in parentheses (clustering at individual level). *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. Strategy method allocations only. In each regression we control for the trustee's ethnicity, age, number of household members, a dummy for literacy, years spent living continuously in Mazar-e-Sharif, log of income (AFN), a dummy for whether the individual has ever signed a contract and an index of perceptions of trust and fairness towards others (3 questions). The F-test compares the sanctioning and no sanctioning condition coefficients.

Table 2.A.5: Investors' Behavior in Games: by Treatment and Response to a World Values Survey Question
 "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?"

Sample	Others cannot be trusted			Others can be trusted			Difference (1) - (4) (T-test p-value)	Difference (2) - (5) (T-test p-value)
	Ingroup (1)	Outgroup (2)	Difference (1) - (2) (T-test p-value) (3)	Ingroup (4)	Outgroup (5)	Difference (4) - (5) (T-test p-value) (6)		
Trust game								
Amount sent	56.60 (24.25)	59.05 (25.28)	-2.45 (0.71)	57.72 (23.38)	55.24 (19.50)	2.48 (0.53)	-1.12 (0.81)	3.81 (0.47)
Share requested back	0.53 (0.22)	0.41 (0.22)	0.11** (0.05)	0.51 (0.22)	0.43 (0.20)	0.08** (0.03)	0.02 (0.71)	-0.01 (0.80)
Belief about share returned	0.51 (0.20)	0.40 (0.17)	0.10** (0.04)	0.52 (0.20)	0.46 (0.16)	0.05 (0.11)	-0.01 (0.80)	-0.06 (0.15)
Share of fair allocations	0.81 (0.40)	0.90 (0.30)	-0.10 (0.33)	0.86 (0.35)	0.94 (0.25)	-0.08 (0.16)	-0.05 (0.49)	-0.03 (0.63)
Requested profit in AFN (investor)	131.91 (40.89)	117.62 (50.98)	14.30 (0.22)	129.65 (44.76)	111.27 (34.71)	18.38*** (0.01)	2.27 (0.79)	6.35 (0.52)
Requested profit in AFN (trustee)	181.28 (52.07)	200.48 (63.76)	-19.20 (0.19)	185.79 (56.50)	199.21 (58.57)	-13.42 (0.21)	-4.51 (0.68)	1.27 (0.93)
Expected profit in AFN (investor)	123.40 (38.18)	115.71 (44.79)	7.69 (0.47)	121.05 (40.21)	109.84 (35.72)	11.21 (0.11)	2.35 (0.76)	5.87 (0.54)
Expected profit in AFN (trustee)	189.79 (58.37)	202.38 (58.81)	-12.59 (0.42)	194.39 (56.25)	200.63 (58.58)	-6.25 (0.55)	-4.60 (0.68)	1.75 (0.91)
Sanctioning game								
Amount sent	53.83 (22.80)	60.00 (22.14)	-6.17 (0.30)	57.72 (21.30)	55.56 (18.82)	2.16 (0.56)	-3.89 (0.37)	4.44 (0.37)
Share requested back	0.54 (0.23)	0.42 (0.20)	0.12** (0.04)	0.55 (0.22)	0.51 (0.19)	0.04 (0.36)	-0.01 (0.85)	-0.09* (0.06)

Continued on next page

Table 2.A.5 – continued from previous page

<i>Sample</i>	<i>Others cannot be trusted</i>			<i>Others can be trusted</i>			Difference (1) - (4) (T-test p-value)	Difference (2) - (5) (T-test p-value)
	<i>Ingroup</i> (1)	<i>Outgroup</i> (2)	Difference (1) - (2) (T-test p-value) (3)	<i>Ingroup</i> (4)	<i>Outgroup</i> (5)	Difference (4) - (5) (T-test p-value) (6)		
Belief about share returned	0.51 (0.20)	0.40 (0.17)	0.10** (0.04)	0.52 (0.20)	0.46 (0.16)	0.05 (0.11)	-0.01 (0.80)	-0.06 (0.15)
Sanction applied	0.28 (0.45)	0.33 (0.48)	-0.06 (0.64)	0.44 (0.50)	0.40 (0.49)	0.04 (0.65)	-0.16* (0.09)	-0.06 (0.61)
Share of fair allocations	0.81 (0.40)	1.00 (0.00)	-0.19** (0.03)	0.77 (0.42)	0.89 (0.32)	-0.12* (0.09)	0.04 (0.65)	0.11 (0.11)
Requested profit in AFN (investor)	132.13 (47.36)	112.38 (39.99)	19.75 (0.10)	137.02 (49.46)	128.10 (34.87)	8.92 (0.25)	-4.89 (0.61)	-15.71* (0.09)
Requested profit in AFN (trustee)	175.53 (56.56)	207.62 (67.67)	-32.09** (0.05)	178.42 (51.09)	183.02 (46.16)	-4.59 (0.61)	-2.89 (0.79)	24.60* (0.07)
Expected profit in AFN (investor)	123.83 (33.07)	112.86 (34.95)	10.97 (0.22)	129.29 (40.40)	119.05 (30.41)	10.24 (0.12)	-5.46 (0.46)	-6.19 (0.44)
Expected profit in AFN (trustee)	183.83 (55.93)	207.14 (54.23)	-23.31 (0.11)	186.43 (53.92)	192.06 (51.59)	-5.63 (0.56)	-2.60 (0.81)	15.08 (0.26)
Dictator game								
Amount sent	40.00 (27.19)	48.10 (24.21)	-8.10 (0.25)	47.72 (25.64)	45.71 (24.74)	2.01 (0.66)	-7.72 (0.14)	2.38 (0.70)
Observations	47	21	68	57	63	120	104	84

Note: Means reported in Columns 1, 2, 4, and 5. Standard deviations in parentheses. Columns 3 and 6 report the difference in means between the *ingroup* and the *outgroup* treatment. Columns 7 and 8 report the difference between columns 1 and 4, and 2 and 5, respectively. P-values of a two-sided t-test are reported in parentheses in columns 3, 6, 7, and 8. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level.

Table 2.A.6: Investors Results Across Dictator, Trust and Sanctioning Games

Sample Dependent variable	Investors							
	Amount sent and requested return				Amount sent and expected return			
	Own profit		Partner's payoff		Own Profit		Partner's payoff	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ingroup	10.54*** (3.68)	18.40*** (6.09)	-11.32 (7.24)	-16.48* (8.50)	7.79** (3.35)	11.20* (5.70)	-8.53 (7.33)	-9.27 (8.54)
Sanctioning game	7.68** (3.86)	11.83** (5.69)	-8.76* (4.62)	-11.34 (6.94)	5.57* (3.22)	6.22 (4.68)	-6.58 (4.30)	-5.73 (5.95)
Ingroup x Sanctioning game		-7.46 (7.74)		4.64 (9.31)		-1.16 (6.45)		-1.53 (8.55)
Dictator game	-68.00*** (3.31)	-59.02*** (4.58)	44.32*** (5.12)	38.29*** (7.16)	-62.59*** (3.12)	-57.56*** (4.60)	38.92*** (5.22)	36.83*** (7.30)
Ingroup x Dictator game		-16.12** (6.49)		10.83 (10.16)		-9.04 (6.24)		3.75 (10.37)
Constant	107.43*** (8.71)	103.06*** (8.94)	119.44*** (18.47)	122.31*** (18.50)	96.07*** (7.22)	94.16*** (7.61)	130.68*** (18.37)	131.10*** (18.14)
Observations	555	555	555	555	554	554	554	554
R-squared	0.47	0.47	0.15	0.15	0.48	0.49	0.13	0.13
F-test p-value								
Sanction(1 + Ingroup) = 0		0.41		0.28		0.26		0.24
Dictator(1 + Ingroup) = 0		0.00		0.00		0.00		0.00

Note: OLS coefficients. Robust standard errors in parentheses. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. The excluded category are *outgroup* trust game results. In each regression we control for the investor's ethnicity, age, number of household members, a dummy for literacy, years spent living continuously in Mazar-e-Sharif, log of income (AFN), a dummy for whether the individual has ever signed a contract and an index of perceptions of trust and fairness towards others (3 questions).

Table 2.A.7: Trustees' Behavior in Games: By Ethnicity and Treatment

Panel A: Tajik subjects			
<i>Sample</i>	<i>Ingroup</i> (1)	<i>Outgroup</i> (2)	Difference (1)-(2) (T-test p-value) (3)
Trust game			
Share returned	58.22 (24.05)	44.08 (21.15)	14.15*** (0.00)
Observations	84	80	
Sanction game			
<i>Sanction applied</i>			
Share returned	60.97 (25.54)	56.44 (24.27)	4.54 (0.25)
Observations	82	78	
<i>Sanction not applied</i>			
Share returned	56.36 (24.43)	47.11 (21.74)	9.25** (0.01)
Observations	86	82	
Panel B: Hazara subjects			
<i>Sample</i>	<i>Ingroup</i> (1)	<i>Outgroup</i> (2)	Difference (1)-(2) (T-test p-value) (3)
Trust game			
Share returned	58.10 (21.90)	42.59 (22.84)	15.50*** (0.00)
Observations	89	99	
Sanction game			
<i>Sanction applied</i>			
Share returned	61.82 (26.14)	60.55 (23.76)	1.27 (0.73)
Observations	88	99	
<i>Sanction not applied</i>			
Share returned	59.28 (25.21)	52.07 (20.90)	7.21** (0.03)
Observations	92	100	

Note: Means reported in Columns 1 and 2. Standard deviations in parentheses. Strategy method allocations only. Column 3 reports the difference in means between the *ingroup* and the *outgroup* treatment. P-values of a two-sided t-test are reported in parentheses in column 3. *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level.

Table 2.A.8: Effect of Sanctions on Share Returned by Ethnicity Across Treatments (OLS)

<i>Sample</i>	<i>Tajik</i>			<i>Hazara</i>		
	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>	<i>All</i>	<i>Ingroup</i>	<i>Outgroup</i>
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)
Ingroup	15.43*** (4.03)			14.37*** (3.57)		
Sanctioning condition	13.79*** (3.90)	-0.64 (3.56)	13.51*** (3.98)	13.97*** (3.16)	0.66 (2.64)	14.28*** (3.29)
Ingroup x Sanctioning condition	-13.89*** (5.17)			-12.79*** (4.09)		
No sanctioning condition	3.26 (3.56)	-1.33 (2.58)	3.06 (3.62)	7.01** (2.67)	0.07 (2.48)	7.42*** (2.70)
Ingroup x No sanctioning condition	-4.18 (4.37)			-6.71* (3.61)		
Amount sent (ths AFN)	-0.10** (0.04)	-0.06 (0.06)	-0.16*** (0.04)	-0.19*** (0.04)	-0.24*** (0.05)	-0.20*** (0.07)
Share requested	0.39*** (0.05)	0.54*** (0.07)	0.29*** (0.06)	0.36*** (0.05)	0.45*** (0.07)	0.28*** (0.06)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22.40*** (7.25)	25.67** (11.57)	24.06*** (8.53)	30.84*** (10.83)	25.84 (18.64)	40.77*** (11.00)
Observations	450	222	228	549	269	280
R-squared	0.30	0.34	0.28	0.28	0.33	0.29
F-test						
H_0 : Sanctioning equals no sanctioning						
<i>Ingroup</i> p-value	0.76	0.81		0.69	0.79	
<i>Outgroup</i> p-value	0.00		0.00	0.03		0.04

Note: OLS coefficients. Standard errors in parentheses (clustering at individual level). *** denotes significance at 1 percent level, ** at 5 percent level and * at 10 percent level. Strategy method allocations only. In each regression we control for the trustee's ethnicity, age, number of household members, a dummy for literacy, years spent living continuously in Mazar-e-Sharif, log of income (AFN), a dummy for whether the individual has ever signed a contract and an index of perceptions of trust and fairness towards others (3 questions). The F-test compares the sanctioning and no sanctioning condition coefficients.

2.A.1 Experiment Instructions

Introduction to the Experiment

Hello, my name is *[experimenter's name]* and this is *[second experimenter's name]*. Thank you for agreeing to participate in this study that concerns the economics of decision making. You will get 100 AFN just for coming, which we will pay you at the end of the session today. Depending on the decisions that you make, and the decisions that others make, you may receive more than this.

You should understand that this is not our own money. This money was given to us by our University for research. This is a one-time payment and will not be repeated in the future. The simulations are part of a scientific study. They will NOT be used to evaluate you or your community. There are no "right" or "wrong" answers. Before we proceed any further, let me stress something that is very important. Many of you were invited here without understanding very much about what we are planning to do today. If at any time you find that this is something that you do not wish to participate in for any reason, you are of course free to leave whether we have started the task or not and the initial fee is yours to keep.

This study may take about three hours, so if you think you will not be able to stay that long without leaving, please let us know now.

I'd like to ask all of you not to talk amongst yourselves from this point on. This is really important and we will have to ask you to leave and you will not have a chance to receive extra money. If you have a question or concern at any time, feel free to ask me or one of my colleagues.

Also, you cannot leave the room during the activities and you cannot use your mobile phone. If you break any of these rules we will have to ask you to leave, and you will receive only your participation fee. Please turn off your mobile phones now.

In order for this study to be carried out correctly, we really need you to not talk about the task while we are here together. This is very important and please be sure that you obey this rule, because it is possible for one person to spoil the task for everyone, in which case we would not be able to continue with the study.

General Introduction to Tasks

This activity will be divided into 3 parts, which we will refer to as "tasks." These decisions will affect the amount of money that you earn. The money that you earn in the tasks is totally separate from the 100 AFN that you will be paid today just for showing up.

However, you won't be paid for all 3 of the tasks: we will draw a number, randomly, at the end of the session to choose one of the 3 tasks and you will be paid for that task only. Since you won't know which task will be paid until after making all decisions, please, think about your decision in each of the three tasks carefully. Also, all of your decisions in all three tasks will be kept private. We will not share information on who made what decisions with anyone outside of the research team.

In this task, you will be matched with a partner. This partner is not here in this room, and you will not know the identity of this person, nor will he or anyone besides us ever know with whom you are matched. This person will come from a similar group like the one today, and was invited to a similar session in the same way that you were invited to the session today. He also received 100 AFN for coming

to a similar session, as you have. Although we can't tell you who this person is, we will tell you a few things about the person you will be matched with. We can tell you that your Partner is a man between the ages of 18-60 years old, who is the head of his household and married with at least one child. He will take part in a session like this one, but in different community where he lives. We will not say the name of this community, but it is a mostly *[Hazara/Tajik]* community that is a similar distance from Mazar-e-Sharif center as this community.

The amount of money that you earn today depends on your decision and your partner's decision as well.

There are a couple more things I'd like to explain.

Firstly, when you make the decision, you will receive these slips of paper that represent bank notes, not real money. However, we will come back two days from today *[payment details]* and pay you in real money. Because the paper slips will be exchanged for real money, you should carefully think before making choices. We will exchange the money and pay you according to what you put in the envelopes – and your partner's decision – so be sure to put all of the paper slips in the envelopes. Your partner will attend a similar session, and will also be paid two days after he attends a session.

Secondly, this decision that you make is private, and no one but us will ever know what you choose. You each received a number when you came in today *[show an example of the id number]*. Hold on to this number, because you will need it to collect your payment.

Investors' Script

Trust Game — Groups

Now we will begin the first task. Please pay close attention, but don't worry if you don't understand the task completely at first: I will go over examples and you will have plenty of time to ask questions.

In this task, you will receive 100 AFN. Your partner will also be given 100 AFN. This is in addition to the 100 that each of you will get for showing up to the session.

You will be given the chance to send part or all of this money to your partner. We will triple everything that you choose to send, so that we add 20 for every 10 that you send, and your partner will receive 30 for every 10 that you send. After receiving this money, your partner will have the chance to send part of this money back to you. He may choose to send, none, some or all of the money back to you.

Your partner will get his 100 AFN regardless of whether you choose to send anything. Also, he may only send back the extra money that he receives – up to three times the amount you send – and cannot send back his initial 100 AFN. When you choose how much to send, you will also have a chance to request how much you would like your partner to send back to you. You can request that he send back to you none of what he receives, all of what he receives or any amount in between. Your partner will be informed of this request, but may choose to send back any amount: he might choose to send back the amount you requested, or he may send back a different amount – either higher or lower.

Your partner has all of the same information about this task as you, so he knows that you had the chance to send money, that what you sent would be tripled, and that he can send some, none or all of it back. He will also know that you know that he would be able to send some money back. He will also

know that you had the chance to request some of this money back, and will know the amount that you requested back.

Please follow along as I demonstrate, explain this in a bit more detail, and give some examples. Please pay close attention, since your payment will be affected by your decision in this task. If you don't understand everything at first, that's fine. You will have a chance to ask questions after the explanation.

We will use red to represent you, and blue to represent your partner *[point to picture]*. So, when we begin, you have got 100 AFN *[show]* and your partner also has 100 *[show]*. Next, you can choose to send something to him. You will make this choice by putting the money that you want to keep in the red envelope that looks like this *[hold up envelope and point to picture]*, and putting the money that you want to send in the blue envelope that looks like this *[show]*. Notice that there is 100 AFN on the blue envelope. This is to remind you that your partner has 100 AFN at the beginning of this task. You will be given 100 AFN in these bank notes.

You could send nothing, 10, 20, 30, 40, 50, 60, 70, 80, 90, or all 100 *[count out bills]*. We will triple anything that you send, so that your partner receives 30 AFN for every 10 you send. He gets nothing if you send 0, he gets 30 if you send 10 and he gets 60 if you send 20; he gets 150 if you send 50; he gets 240 if you send 80; and 300 if you send all 100.

Next, we will collect those envelopes. And you will have the chance to decide how much you'd like to request back. Let's say that you choose to send 10 AFN. *[show]*. You put 90 AFN in the red envelope to keep, and 10 AFN in the blue envelope to send to your partner. Remember anything that you send gets tripled, so we add 20 AFN to the 10 AFN in the blue envelope *[show]*. The money you keep stays the same. *[show]*.

Okay, now your partner receives 30 and now has this, in addition to the 100 he had at the beginning of the task, so 130 in total *[show]*. You're left with 90. *[show]*. You may request anywhere from 0 to 30 AFN back from your partner. He will be provided with this information, but will have the choice to send back any amount he chooses.

Let's say you request 20 back. If he sends back this 20, then you will get 110 — that is, the 90 you kept plus the 20 that he sends back. Your partner will get the 100 he started out with, plus the 30 he receives (the 10 you sent, tripled) minus the 20 he sent back, which is 110 in total.

If your partner sends back nothing, then you will end this task with 90 *[show]*, and your partner will end up with 130 *[show]* —that's the 100 he started the experiment with, plus the 30 he received.

Further examples given.

Any questions? *[take and answer questions]*.

Now, we will invite you one by one and ask you a couple of questions to make sure that you understand the task.

Control questions:

1. Which envelope is for keeping? Sending?
2. Which envelope will be tripled?

Trust Game — Individually

Now we will begin the task. We will explain how you will make your decision in more detail, as we go along. Again, feel free to ask any questions if there is something that you do not understand.

Here is the 100 AFN in our banknotes. Remember the red envelope is for the money that you want to keep, and what you put in the blue envelope will be sent to your partner. Your partner will receive triple this amount, that means that we add 20 AFN to every 10 AFN you send, and will be able to return a portion of what he receives. Also remember that you will have the chance to request an amount back. You must put all 100 in the envelopes!

We will tell your partner some information about you: that you're a man between 18-60 years old, who is the head of your household and married with at least one child. That you took part in a session as he did, but in a different community than him – the community where you live. We will not tell him the name of your community, but will tell him that that is a mostly *[Hazara/Tajik]* community and that it's a similar distance from Mazar-e-Sharif center as his community.

Now we will tell you about your partner: Your partner is a man between the ages of 18-60 years old, who is the head of his household and married has at least one child. He will take part in a session like this one, but in different community where he lives. We will not say the name of this community, but it is a mostly *[Hazara/Tajik]* community that is a similar distance from Mazar-e-Sharif center as this community.

Please make your decision now.

[experimenter: collect envelopes]

Please wait while we collect the envelopes and count the results.

Now, you will decide how much to request back from your partner. You have already made your decision, and the amount you request back will not change this decision. Your partner will be informed of the amount that you requested back. He may choose to send back this amount, or may decide to send back any different amount. Here is the new set of bank notes. *[enumerators: distribute the banknotes]*. The total amount we will give you is equal to the amount you sent your partner, tripled. This is the amount your partner will receive from you. Remember that you already have the money you kept from the previous decision. That is represented by this red card that we have given you *[show card]*. Also, remember that your partner began the task with 100 AFN, and cannot send this back. To remind you, that money is represented by 100 AFN on the blue card *[show card]*.

You will now decide how much to request back from your partner by putting the money onto the cards as I will now explain. Place the money that you would like to request onto the red card. Your partner will be informed that this is the amount you are requesting back. Leave the remaining banknotes on top of the blue card. This is the money you are requesting your partner to keep. Your partner will be informed that this is the amount you are requesting for him to keep. Remember that all the paper banknotes that you are given must be divided between the two cards. We will now record your requests, so that we can tell your partners.

Next, we would like you to tell us how much you think that your partner will return to you. Whatever you say, it will have no effect on how much you have actually sent. Before, you requested an amount for your partner to return, and we told you that we will inform him of this request. This next decision, however, will NOT be communicated to your partner. No one else, besides us will be informed of this decision. If you do correctly identify the amount sent, you will get an extra 20. This ends the first task.

Now we will move on to the second task.

Sanction Game — Groups

Now we will begin task two. To remind you, you won't be paid for all 3 of the tasks: we will draw a number, randomly, from a bag at the end of the session to choose one of the 3 tasks and you will be paid for that task only. Since you won't know which task will be paid until after making all decisions, please, make your decisions in all tasks carefully. In task two you will be matched with a different partner. This person took part in the same session as your partner from the first task from the same community. However, it's a different individual who will not be informed of the decisions you made in the previous task. In task two you will be matched with the same partner as in the previous task.

We will tell your partner some information about you: that you're a man between 18-60 years old, who is the head of your household and married with at least one child. That you took part in a session as he did, but in a different community than him – the community where you live. We will not tell him the name of your community, but will tell him that that is a mostly [*Hazara/Tajik*] community and that it's a similar distance from Mazar-e-Sharif center as his community.

Now we will tell you about your partner: Your partner is a man between the ages of 18-60 years old, who is the head of his household and married has at least one child. He will take part in a session like this one, but in different community where he lives. We will not say the name of this community, but it is a mostly [*Hazara/Tajik*] community that is a similar distance from Mazar-e-Sharif center as this community.

The task is similar in some ways to the previous task, but please pay very close attention because there are some important differences.

In this task, both you and your partner will again be given 100 AFN to start the task. As in the previous task, you will be able to transfer some, none or all of this 100 AFN to your partner. This is in addition to the 100 that each of you will get for coming. Whatever amount that you choose to send to him will be tripled.

After receiving this money, your partner will have the chance to send part of this money back to you. He may choose to send, none, some or all of the money back to you.

Your partner will keep his 100 AFN regardless of whether you choose to send anything and regardless of the amount he chooses to send back to you. Also, he may only send back the extra money that he receives – up to three times the amount you send – and cannot send back his initial 100 AFN.

When you choose how much to send, you will also have a chance to request how much you would like your partner to send back to you – as in the previous task. You can request that he send back to you none of what he receives, all of what he receives or any amount in between. Your partner will be informed of this request, but may choose to send back any amount: he might choose to send back the amount you requested, or he may send back a different amount – either higher or lower.

There is one important difference from the previous task: in this task you have one more decision to make, which you did not make in the previous task. You can choose whether or not 40 AFN will be deducted from the money your partner will be paid if he sends back an amount less than the amount that you requested.

If you choose NOT to have 40 AFN deducted from your partner, then regardless of what he chooses

to send back to you, he will end the task with his initial 100 AFN, plus the amount you send, tripled, minus the amount he chooses to send back to you.

If you do choose to have 40 AFN deducted from your partner if he returns less than the amount that you request, and he sends back what you have requested or more, then he will again end up with his initial 100 AFN, plus the amount you send, tripled, minus the amount he chooses to send back to you.

If you do choose to have 40 AFN deducted from your partner if he returns less than the amount that you request, and then he sends back less than what you have requested, then he will end up with his initial 100 AFN, plus the amount you send, tripled, minus the amount he chooses to send back to you, minus the 40 AFN deduction.

Your partner has all of the same information about this task as you, so he knows that you had the chance to send money, that what you sent would be tripled, and that he can send some, none or all of it back. He will also know that you know that he would be able to send some money back. He will also know that you had the chance to request some of this money back, and will know the amount that you requested back. He will also know that you had the chance to choose whether 40 AFN would be deducted if he sends back less than the amount that you requested, and whether or not you chose to do so.

Please follow along as I explain this in a bit more detail, and give some examples. As before, please pay close attention, since your payment will be affected by your decision in this task. If you don't understand everything at first, that's fine. You will have a chance to ask questions after the explanation.

To remind you, red represents you *[show clipboard]*, and blue represents your partner *[show clipboard]*. So, when we begin, you have got 100 AFN *[show banknotes]* and your partner also has 100 *[show]*. Next, you can choose to send something to him. You will make this choice by putting the money that you want to keep in the red envelope *[show]*, and putting the money that you want to send in the blue envelope that looks like this *[show]*. Remember that there is 100 AFN on the blue envelope to remind you that your partner has 100 AFN at the beginning of this task *[show]*.

As in the previous task, you can send nothing, all 100 or any amount in between. We will triple anything that you send, so that your partner receives 30 AFN for every 10 you send – just as in the previous task.

Next, you will have the chance to decide how much you'd like to request back. You may request anywhere from 0 to the whole amount that your partner receives – the tripled amount that you send. He will be provided with this information, but will have the choice to send back any amount he chooses.

In this task you have one more decision to make that you did not have in the previous task. You can choose whether or not 40 AFN will be deducted from your partner's earnings if he sends back less than the amount that you requested. You will do this by placing this card *[hold up card]* either on the green side *[show green side]* to indicate that 40 AFN will be deducted from your partner's earnings if he sends back less than the amount you requested, or on the yellow side *[hold up yellow side]* which indicates that no money will be deducted if your partner returns less money than you requested.

Now, we will go over an example to make this clearer. You and your partner begin this task with 100 AFN *[show]*. Let's say that you decide to send 10 AFN to your partner *[show]*. Now you have 90 AFN. The 10 AFN that you send is tripled, so that he receives 30 AFN *[show]*. Together with the initial 100 AFN, he now has 130. You can request that he send back anywhere from 0 to 30 AFN. Let's say

that you request 20 back.

Next, you have the choice of whether he will have 40 AFN deducted from his earnings if he sends back less than 20. Let's say that you choose to deduct 40 AFN from your partner's earnings if he chooses to send back less than the amount that you request. You will place the card down with the green side facing up *[show]*.

Now, since you have requested 20 back from your partner, if he sends back 20, then you will get 110 – that is, the 90 you kept plus the 20 that he sends back. Your partner will get the 100 he started out with, plus the 30 he receives (the 10 you sent, tripled) minus the 20 he sent back, which is 110 in total. You chose to have 40 AFN deducted from your partner's payoff if he sent back less than the amount you requested. Since he sent back the amount you requested, then 40 AFN will not be deducted from his earnings.

Further examples given.

Any questions? *[take and answer questions]*.

Sanction Game — Individually

Now we will begin the task. Again, feel free to ask any questions if there is something that you do not understand. Here are 100 AFN in our banknotes. Remember the red envelope is for the money that you want to keep, and what you put in the blue envelope will be sent to your partner. Your partner will receive triple this amount, that means that we add 20 AFN to every 10 AFN you send, and will be able to return a portion of what he receives. Also remember that you will have the chance to request an amount back, and to decide whether 40 AFN will be deducted from your partner's earnings if he sends back less than this amount. You must put all 100 in the envelopes! Please make your decision now.

We will tell your partner some information about you: that you're a man between 18-60 years old, who is the head of your household and married with at least one child. That you took part in a session as he did, but in a different community than him – the community where you live. We will not tell him the name of your community, but will tell him that it is a mostly *[Hazara/Tajik]* community and that it's a similar distance from Mazar-e-Sharif center as his community.

Now we will tell you about your partner: Your partner is a man between the ages of 18-60 years old, who is the head of his household and married has at least one child. He will take part in a session like this one, but in different community where he lives. We will not say the name of this community, but it is a mostly *[Hazara/Tajik]* community that is a similar distance from Mazar-e-Sharif center as this community.

[experimenter: collect envelopes]

Please wait while we collect the envelopes and count the results.

Now, you will decide how much to request back from your partner. You have already made your decision, and the amount you request back will not change this decision. Your partner will be informed of the amount that you requested back. He may choose to send back this amount, or may decide to send back any different amount. Now you will also decide whether 40 AFN will be deducted from your partner if he returns less than the amount you request. You will do this by placing this card with the green side facing up to indicate that 40 AFN will be deducted if he returns less than you request. If you place the card with the yellow side facing up, then nothing will be deducted from your partner if he

sends less than you have requested.

Here is a new set of bank notes. The total amount we will give you is equal to the amount you sent your partner, tripled. This is the amount your partner will receive from you. Remember that you already have the money you kept from the previous decision. That is represented by this red card that we have given you *[show card]*. Also, remember that your partner began the task with 100 AFN, and cannot send this back. To remind you, that money is represented by 100 AFN note on the blue card *[show card]*.

You will now decide how much to request back from your partner by putting the money onto the cards as I will now explain. Place the money that you would like to request onto the red card. Your partner will be informed that this is the amount you are requesting back. Leave the remaining banknotes on top of the blue card. This is the money you are requesting your partner to keep. Your partner will be informed that this is the amount you are requesting for him to keep. Remember that all the paper banknotes that you are given must be divided between the two cards. We will now record your requests, so that we can tell your partners.

Next, we would like you to tell us how much you think that your partner will return to you. Whatever you say, it will have no effect on how much you have actually sent. Before, you requested an amount for your partner to return, and we told you that we will inform him of this request. This next decision, however, will NOT be communicated to your partner. No one else, besides us. If you do correctly identify the amount sent, you will get an extra 20 AFN.

This ends the second task. Now we will move on to the third task.

Receivers' Script

Trust Game — Groups

Now we will begin the first task. Please pay close attention, but don't worry if you don't understand the task completely at first: I will go over examples and you will have plenty of time to ask questions.

Today we will use these slips of paper to represent banknotes *[show]*. We will exchange these slips of paper for real money and come back *[time of payment]* to pay you your earnings in cash (AFN).

In this task, you will receive 100 AFN. Your partner will also be given 100 AFN. This is in addition to the 100 that each of you will get for showing up today.

Your partner will be given the chance to send part or all of this money to you. We will triple everything that he chooses to send, so that we add 20 for every 10 that he sends, so that you will receive 30 for every 10 that he sends. After receiving this money, you will have the chance to send part of this money back to him. You may choose to send none, some or all of the money back to him.

You will receive the 100 AFN regardless of whether your partner chooses to send anything to you. Also, you may only send back the extra money that you receive – up to three times the amount your partner sends – and cannot send back your initial 100 AFN.

When your partner chooses how much to send, he will also have a chance to request how much he would like you to send back to him. He can request that you send back none of what you received, all of what you received or any amount in between. You will be informed of this request, but may choose to send back any amount: you might choose to send back the amount requested, or you may send back a

different amount – either higher or lower.

Your partner has all of the same information about this task as you, so he knows that he could send money, that what he sent would be tripled, and that you could send some, none or all of it back. He also knew that you would be informed of the amount that he requested back from you.

Please follow along as I explain this in a bit more detail, and give some examples. Please pay close attention, since your payment will be affected by your decision in this task. If you don't understand everything at first, that's fine. You will have a chance to ask questions after the explanation.

We will use blue to represent you, and red to represent your partner *[show clipboards]*. So, when we begin, you have got 100 AFN and your partner also has 100 as you can see here *[show envelopes and banknotes]*. Next, he can choose to send something to you.

He could send nothing, 10, 20, 30, 40, 50, 60, 70, 80, 90, or all 100 *[count out bills]*. We will triple anything that he sends, so that you receive 30 AFN for every 10 he sends. If he sends 0, you won't get anything, if he sends 10, you will get 30, if he sends 20 you will get 60; if he sends 50 he gets 150, if he sends 80 you will get 240, and if he sends all 100, you will get 300.

Let's say that your partner chooses to send 10 AFN. *[show]*. Remember anything that he sends gets tripled, so we add 20 AFN to the 10 AFN, and you will receive 30. The money that your partner keeps stays the same. *[show]*.

Okay, now you receive 30, and in addition to the 100 that you had at the beginning of the task, you now have 130 in total. Your partner is left with 90 *[show]*. Your partner could request anywhere from 0 to 30 AFN back from you. You will be provided with this request, but will have the choice to send back any amount that you choose.

Let's say that he requests 20 back. You will be informed of this amount in the following way. See this card with the red stripe *[show]*? This represents the money that your partner kept. Since he sent 10, he kept 90. So there is 90 AFN on the card. The blue card contains the amount he sent, which was tripled here *[point to banknotes]*, in addition to the 100 AFN that you started with. Your partner can request any of the tripled amount back, but not this 100. We will place the banknotes on the blue and red cards according the amounts that your partner requested that you keep and requested that you send back. Remember, you can choose to send back this amount or a different amount, either lower or higher than the requested amount.

You will make your decision by placing the banknotes into the envelope. You place the banknotes that you wish to keep in the blue envelope and leave it on the blue card. You will place the banknotes you want to send back to your partner in the red envelope and put it on the red card. You must put all of the bank notes in one of the two envelopes.

If you send back 20, then your partner will get 110 – that is, the 90 you kept plus the 20 that he sends back *[show]*. You will get the 100 that you start out with, plus the 30 you receive from your partner (the 10 he sent, tripled) minus the 20 he sent back, which is 110 in total *[show]*. Additional examples given.

There is one more thing that we need to explain: when you do the task, you will decide how much to return a few times. Each time, you will be given an amount sent, and amount requested back from your partner. However, only one these will be from your actual partner, and you will be paid for this decision only. You will not know ahead of time which decision is actually what your partner decided, and therefore, you should make each decision as if it were the real one. The person who will actually be

your partner will only be paid for the decision you make regarding the amount that he actually sent as well.

Trust Game — Individually

Now we will begin the task. We will explain how you will make your decision in more detail, as we go along. Again, feel free to ask any questions if there is something that you do not understand.

Now is the first decision you will make, based on the potential decisions of your partner. Here is the money that is sent, (amount sent), and we then triple that amount. So you get (amount tripled). You started the experiment with 100, which you can see here, so together you've got (total amount). Your partner now has (amount partner kept) which you can see here on the red card. Now, your partner requested (amount requested) back. So, I will put (amount requested) back on the red card to represent this. Your partner requested that you keep (amount requested keep), so I will put that money on the blue card to represent this. That means he would have (amount requested plus amount kept) and you would have (amount requested keep plus 100). Remember, you may choose to keep and send back these amounts, or you can choose to keep and send back any different amount that you wish. You will make your decision by taking the banknotes and putting them into this blue envelope to keep, and the red envelope to send back to the partner. You must put all of the bank notes in one envelope or the other.

OK, now I will leave you to make your decision.

Remember, we told your partner some information about you: We told him the same information about you: that you're a man between 18-60 years old, who is the head of your household and married with at least one child. That you would take part in a session as he did, but in a different community than him – the community where you live. We did not tell him the name of your community, but did tell him that it is a mostly *[Hazara/Tajik]* community and that it's a similar distance from Mazar-e-Sharif center as his community.

Now we will tell you about your partner: Your partner is a man between the ages of 18-60 years old, who is the head of his household, married and has at least one child. He took part in a session like this one, but in different community where he lives. We will not say the name of this community, but it is a mostly *[Hazara/Tajik]* community that is a similar distance from Mazar-e-Sharif center as this community.

Sanction Game — Groups

Now we will begin task two. To remind you, you won't be paid for all 3 of the tasks: we will draw a number, randomly, from a bag at the end of the session to choose one of the 3 tasks and you will be paid for that task only. Since you won't know which task will be paid until after making all decisions, please, make your decisions in all tasks carefully. In task two you will be matched with a different partner. This person took part in the same session as your partner from the first task from the same community. However, it's a different individual who will not be informed of the decisions you made in the previous task.

We told your partner some information about you: We told him the same information about you: that you're a man between 18-60 years old, who is the head of your household and married with at least one child. That you would take part in a session as he did, but in a different community than him – the

community where you live. We did not tell him the name of your community, but did tell him that it is a mostly [Hazara/Tajik] community and that it's a similar distance from Mazar-e-Sharif center as his community.

Now we will tell you about your partner: Your partner is a man between the ages of 18-60 years old, who is the head of his household, married and has at least one child. He took part in a session like this one, but in different community where he lives. We will not say the name of this community, but it is a mostly [Hazara/Tajik] community that is a similar distance from Mazar-e-Sharif center as this community.

The task is similar in some ways to the previous task, but please pay very close attention because there are some important differences.

In this task, you will receive 100 AFN. Your partner will also be given 100 AFN. This is in addition to the 100 that each of you will get for showing up today.

As in the previous task, your partner will be given the chance to send part or all of this money to you. Whatever amount that you choose to send to him will be tripled.

After receiving this money, you will have the chance to send part of this money back to him. You may choose to send none, some or all of the money back to him.

You will receive the 100 AFN regardless of whether your partner chooses to send anything to you. Also, you may only send back the extra money that you receive – up to three times the amount your partner sends – and cannot send back your initial 100 AFN.

When your partner chooses how much to send, he will also have a chance to request how much he would like you to send back to him. He can request that you send back none of what you received, all of what you received or any amount in between. You will be informed of this request, but may choose to send back any amount: you might choose to send back the amount requested, or you may send back a different amount – either higher or lower.

There is one important difference from the previous task: Your partner had one more choice to make. He chose whether or not 40 AFN would be deducted from the money you will be paid if you send back an amount less than the amount that he requested.

If he chooses NOT to have 40 AFN deducted from your earnings if you send back less than what is requested, then regardless of what you choose to send back to him, you will end the task with the initial 100 AFN, plus the amount your partner sent to you, tripled, minus the amount you choose to send back to him.

If your partner does choose to have 40 AFN deducted from your earnings if you return less than the amount that he requested, and you send back what he has requested or more, then you will again end up with the initial 100 AFN, plus the amount he sent to you, tripled, minus the amount you choose to send back to him.

If your partner does choose to have 40 AFN deducted from your earnings if you return less than the amount that he requested, and then you send back less than what he has requested, then you will end up with the initial 100 AFN, plus the amount you he sent to you, tripled, minus the amount he chooses to send back to you, minus the 40 AFN deduction.

Your partner has all of the same information about this task as you, so he knows that he could send money, that what he sent would be tripled, and that you could send some, none or all of it back. He also

knew that you would be informed of the amount that he requested back from you, and that you would be informed that he could choose whether 40 AFN would be deducted from your earnings if you sent back less than he requested, and whether or not he chose to do so.

Please follow along as I explain this in a bit more detail, and give some examples. As before, please pay close attention, since your payment will be affected by your decision in this task. If you don't understand everything at first, that's fine. You will have a chance to ask questions after the explanation.

To remind you, blue represents you *[show clipboard]*, and red represents your partner *[show clipboard]*. So, when we begin, you have got 100 AFN *[show banknotes]* and your partner also has 100 *[show]*. Next, your partner can choose to send something to you.

As in the previous task, he can send nothing, all 100 or any amount in between. We will triple anything that he sends, so that you will receive 30 AFN for every 10 you send – just as in the previous task.

Let's say that your partner chooses to send 10 AFN. Okay, now you receive 30 and now have this, in addition to the 100 you had at the beginning of the task, so 130 in total. Your partner is left with 90. *[point to picture]*. Your partner could request anywhere from 0 to 30 AFN back from you. You will be provided with this information, but will have the choice to send back any amount you choose, just as in the last task.

Your partner could also choose whether or not 40 AFN will be deducted from your earnings if you send back less than the amount that he requested. We will communicate whether or not he chose to do this by placing this card either green-side-up or yellow-side-up. Green indicates that your partner has chosen to deduct 40 AFN from you. Yellow indicates that he has not chosen to deduct anything if you send back less than the amount that he has requested. Let's say that he chooses to have 40 AFN deducted from your earnings if you send back less than the amount that you requested. So you place this card with the green side facing up.

Let's say he requests 20 back. If you send back this 20, then you will get the 100 you started out with, plus the 30 you receive (the 10 you sent, tripled) minus the 20 you sent back, which is 110 in total. Your partner will get 110 as well, since he started with 100, sent 10 to you, then received 20 back from you.

He requested 20 back, and chose to have 40 AFN deducted from your partner's earnings if he sends back less than the amount you requested *[show green/yellow card]*. Since you sent back the amount that you requested, this 40 AFN will not be deducted.

Further examples given.

There is one more thing that we need to explain: when you do the task, you will decide how much to return a few times, just as you did last time. Each time, you will be given an amount sent, and amount requested back from your partner and whether your partner decided to have 40 AFN deducted from your payoff if you send back less than the amount requested. However, only one these will be from your actual partner, and you will be paid for this decision only. You will not know ahead of time which decision is actually what your partner decided, and therefore, you should make each decision as if it were the real one. The person who will actually be your partner will only be paid for decision you make regarding the amount that he actually sent as well.

Any questions? *[take and answer questions]*.

Now, we will invite you one by one and ask you a couple of questions to make sure that you understand the task.

Sanction Game — Individually

Now we will begin the task. We will explain how you will make your decision in more detail, as we go along. Again, feel free to ask any questions if there is something that you do not understand.

Now is the first decision you will make, based on the potential decisions of your partner. Here is the money that is sent, (amount sent), and we then triple that amount. So you get (amount tripled). You started the experiment with 100, which you can see here, so together you've got (total amount). Your partner now has (amount partner kept) which you can see here on the red card. Now, your partner requested (amount requested) back. So, I will put (amount requested) back on the red card to represent this. Your partner requested that you keep (amount requested keep), so I will put that money on the blue card to represent this. That means he would have (amount requested plus amount kept) and you would have (amount requested keep plus 100). Remember, you may choose to keep and send back these amounts, or you can choose to keep and send back any different amount that you wish. You will make your decision by taking the banknotes and putting them into this blue envelope to keep, and the red envelope to send back to the partner. You must put all of the bank notes in one envelope or the other.

OK, now I will leave you to make your decision.

We told your partner some information about you: We told him the same information about you: that you're a man between 18-60 years old, who is the head of your household and married with at least one child. That you would take part in a session as he did, but in a different community than him – the community where you live. We did not tell him the name of your community, but did tell him that it is a mostly [*Hazara/Tajik*] community and that it's a similar distance from Mazar-e-Sharif center as his community.

Now we will tell you about your partner: Your partner is a man between the ages of 18-60 years old, who is the head of his household, married and has at least one child. He took part in a session like this one, but in different community where he lives. We will not say the name of this community, but it is a mostly [*Hazara/Tajik*] community that is a similar distance from Mazar-e-Sharif center as this community.

Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition

Vojtěch Bartoš, Michal Bauer, Julie Chytilová, and Filip Matějka¹

Abstract

We integrate tools to monitor information acquisition in field experiments on discrimination and examine whether gaps arise already when decision-makers choose the effort level for reading an application. In both countries we study, negatively stereotyped minority names reduce employers' effort to inspect resumes. In contrast, minority names increase information acquisition in the rental housing market. Both results are consistent with a model of endogenous allocation of costly attention, which magnifies the role of prior beliefs and preferences beyond the one considered in standard models of discrimination. The findings have implications for magnitude of discrimination, returns to human

¹This work is forthcoming in the *American Economic Review*. We thank Colin Camerer, Stefano DellaVigna, Randy Filer, Martin Gregor, Christian Hellwig, Štěpán Jurajda, Peter Katuščák, Ulrike Malmendier, Marti Mestieri, Sendhil Mullainathan, Ron Oaxaca, Franck Portier, Chris Sims, Jakub Steiner, Matthias Sutter, five anonymous referees and audiences at Oxford University, NYU, Princeton University, Columbia University, CERGE-EI, Toulouse School of Economics, the University of Gothenburg, the Institute for Advanced Studies in Vienna, UC Berkeley, University of San Francisco, and ASSA meeting in Philadelphia for valuable comments, and Kateřina Boušková, Lydia Hähnel, Vít Hradil, Iva Pejsarová, Lenka Švejdrová and Viktor Zeisel for excellent research assistance. Institutional Review Board approval has not been obtained because the institutions, which the authors are affiliated with, do not have IRBs. This research was supported by a grant from the CERGE-EI Foundation under a program of the Global Development Network and by a grant from the Czech Science Foundation (13-20217S).

capital and policy.

3.1 Introduction

Understanding why people discriminate based on ethnicity, gender, or other observable group attributes has been one of the central topics in economics and other social sciences for decades.² Since the seminal work of Phelps (1972) and Arrow (1973), it has been widely acknowledged that due to a lack of individual-level information decision makers often rely on a group attribute as a signal of unobserved individual characteristics. This may give rise to "statistical discrimination" in selection decisions on various markets.³ At the same time, a large body of research in both economics and psychology documents that scarce attention plays an important role in decision making (e.g., Newell, Shaw, and Simon 1958, Kahneman 1973, Gabaix, Laibson, and Moloche 2006, Chetty, Looney, and Kroft 2009, Fehr and Rangel 2011) and theories assuming costly attention made progress in explaining a range of important economic phenomena (e.g., Sims 2003, Mackowiak and Wiederholt 2009).

While the existing models of discrimination implicitly assume that individuals are fully attentive to available information, we link the literatures on discrimination and scarce attention. We develop a model in which we describe how knowledge of a group attribute impacts the level of attention to information about an individual and how the resulting asymmetry in acquired information across groups—denoted "attention discrimination"—can lead to discrimination in a selection decision. To test the model, we build on the experimental design of Bertrand and Mullainathan (2004) and perform three correspondence tests in two countries. A novel feature of our field experiments are the tools to measure the process of decision-making, in addition to selection choices, by monitoring acquisition of information about applicants.

Attention to available information about candidates is crucial input in virtually any selection process: in the recruitment of employees, school admissions, housing market tenant selection, loan provisions, voting in elections, or scientific review processes, to

²Researchers have produced a vast amount of evidence documenting discriminatory behavior based on ethnicity or gender on labor, housing, and consumer markets. Yinger (1998) and Altonji and Blank (1999) survey regression-based (non-experimental) evidence, Riach and Rich (2002) and List and Rasul (2011) summarize related field experiments.

³Taste-based discrimination is the second prominent explanation for why people discriminate (Becker 1971). It arises due to preferences, not due to lack of information.

name a few examples. The Economist (2012), for instance, describes the process as follows:

"They [human resource staff] look at a CV for ten seconds and then decide whether or not to continue reading. If they do, they read for another 20 seconds, before deciding again whether to press on, until there is either enough interest to justify an interview or to toss you into the 'no' pile." ⁴

Similarly, qualitative studies of college admissions describe the reading of applications by admission committees as very coarse and quick (Stevens 2009; Deresiewicz 2014). The pioneering field experiment on discrimination in the labor market by Bertrand and Mullainathan (2004) finds that returns to sending higher-quality resumes, in terms of callbacks, are higher for applicants with a White-sounding name compared to applicants with an African-American-sounding name in the US labor market. The pattern is consistent with employers not continuing to read once they see an African-American name on a resume, thus resulting in greater discrimination among more qualified applicants. These provocative findings motivate the need to find a way to measure the effect of a name on reading effort. For a theory, the findings open the question as to whether choices about inspecting applicants are guided by the expected benefits of reading, as indicated by the qualitative description from practitioners.

To illustrate how the allocation of costly attention affects discrimination, we propose a new model. First, acquiring information is costly and decision makers optimize how much information to acquire based on expected net benefits. This leads to "attention discrimination". Second, imperfect information affects selection decisions because the less the decision maker knows about an individual, the more he relies on observable group attributes when assessing individual quality. Putting these two key features together, the endogenous attention magnifies (in most types of markets) the impact of animus and prior beliefs about group quality. Discrimination in selection decisions can persist even if perfect information about an individual is readily available, if it is equally difficult to screen individuals from dissimilar groups and if there are no differences in preferences. It also implies lower returns to employment qualifications for negatively stereotyped groups

⁴Also, a recent study found that human resource managers spend on average six seconds reviewing an individual resume (TheLadders 2012). Another study (Dechief and Oreopoulos 2012) quotes several recruiters describing the need to have quick routines for selecting resumes: "I'm down to about seven seconds. [The information I'm looking for] needs to pop out so I'm very much onto keyword skimming. I'm almost like a Googlebot, like when you put in a search query. I have to do it really fast. I don't have time to waste. ... I do realize how unfair the whole process is."

in selective markets, and for policy the important role of the timing of when a group attribute is revealed.

The model provides the following testable prediction: In "cherry-picking" markets where only top applicants are selected from a large pool of candidates (e.g., much of the labor market, admissions to top schools, the scientific review process in leading scholarly journals), decision makers should favor acquiring information about individuals from an a priori more attractive group. In contrast, in "lemon-dropping" markets where most applicants are selected (e.g., the rental housing market, admissions to nearly open-access schools), decision makers benefit more from acquiring information about individuals from a less attractive group. This is because more information should be acquired when its expected benefits are higher, which is when there is a higher chance that the informed decision differs from the status quo, i.e. when there is a higher chance of accepting the applicant in a market where most applicants are rejected and vice versa.⁵

We test the predictions of the model by monitoring information acquisition in three field experiments—in rental housing and labor markets in the Czech Republic and in the labor market in Germany. We send emails responding to apartment rental advertisements and to job openings. In each country we study discrimination against negatively stereotyped ethnic minorities and randomly vary the names of fictitious applicants. In the German labor market we also vary the quality of applicants by signaling recent unemployment in the email. To monitor information acquisition in the labor market, employers receive an email application for a job opening, which contains a hyperlink to a resume. Similarly, in the housing market landlords can click on a hyperlink located in the email and learn more on an applicant's personal website. We monitor whether employers and landlords open the applicant's resume (resp. website) as well as the intensity of information acquisition.

While we find strong evidence of discrimination against minorities in a selection decision (invitation to a next stage) on both the housing and labor markets, we also document that systematic discrimination arises even earlier, during the process of information acquisition. The key findings on attention allocation are as follows. In the labor markets in both countries, employers put more effort to opening and reading resumes of majority compared to minority candidates, while on the rental housing market landlords acquire

⁵Bose and Lang (2013) use similar logic in a different setting by showing that costly monitoring in the workplace is most beneficial when the employer has neither too low nor too high priors about the quality of a worker.

more information about minority compared to majority candidates. Signaling unemployment lowers attention to an applicant’s resume, similarly as minority name does. The set of results on attention allocation is consistent with the proposed model of discrimination with endogenous attention. The labor markets we study are very selective, as indicated by low invitation rates, and decision makers acquire less information about a priori less attractive applicants, whether it be a person with minority ethnic status or unemployed. In contrast, the rental housing market is not selective and decision-makers acquire more information about applicants who look a priori less attractive. Later, we also discuss alternative explanations.

Methodologically, our paper contributes to efforts to test theory with enhanced measurement tools. In the lab, researchers have fruitfully complemented choice data with measures of the decision-making process to sort through alternative theoretical explanations of observed behavior. These techniques involve eye-tracking (Knoepfle, Wang, and Camerer 2009; Krajbich, Armel, and Rangel 2010; Arieli, Ben-Ami, and Rubinstein 2011; Reutskaja et al. 2011; Devetag, Di Guida, and Polonio 2016), mouse-tracking⁶ (Camerer et al. 1993; Costa-Gomes, Crawford, and Broseta 2001; Costa-Gomes and Crawford 2006; Gabaix, Laibson, and Moloche 2006; Brocas et al. 2014) or monitoring brain activity (Bhatt and Camerer 2005; Hare, Malmaud, and Rangel 2011). Camerer and Johnson (2004) and Crawford (2008) summarize how progress in testing theories of human behavior has been facilitated by using information acquisition measures. To the best of our knowledge, ours is the first study that integrates monitoring information acquisition, in addition to selection decisions, into a field experiment.

In order to identify discrimination based on ethnicity, gender, caste, or sexual orientation in the labor and housing markets, previous correspondence tests⁷ estimated the effects of a group-attribute signal (mostly names) in applications (e.g., Neumark, Bank, and van Nort 1996, Weichselbaumer 2003, Bertrand and Mullainathan 2004, Ahmed and Hammarstedt 2008, Banerjee et al. 2009, Kaas and Manger 2012). These experiments measure the likelihood of a callback (or invitation) as the outcome of interest.⁸ We of-

⁶Mouse-tracking, a technique closest to the monitoring tools used in this paper, typically uses Mouselab software, which displays information hidden in boxes on the computer screen and then tracks which and how many pieces of information subjects acquire.

⁷Two types of procedures have been used to test for the extent of discrimination on markets (Riach and Rich 2002). Correspondence tests involve responding to vacancies with written applications. Personal approaches, typically referred to as audit tests, include studies that have trained individuals attending job interviews or applying over the telephone.

⁸An important exception is Milkman, Akinola, and Chugh (2012) who study race and gender discrimination in academia and measure not only callback of faculty members reacting to students’ requests to

fer an extension of this widely-used design by measuring effort expended to open and read resumes in the labor market and to acquire information about potential tenants in the rental housing market.⁹ Although the interview invitation decision can also be interpreted as a choice about costly information acquisition, the richer data about the decision-making process is useful for at least two reasons. First, since the costs of reading a resume are tiny compared to interviewing an applicant, it is an open question as to whether discrimination manifests itself already at the very outset of the decision-making process. This is potentially important for policy, since very early signals would have a larger impact on outcomes and also because addressing the smallest frictions in the early stage, such as the cost of reading a resume as opposed to the cost of an interview, might be easier. Second, measures of reading effort allow for a richer understanding of how attention discrimination operates on different types of markets.¹⁰

Our model of attention discrimination contributes to existing theories of discrimination (for a recent survey see Lang and Lehmann 2012). It is related most closely to "screening discrimination" (Cornell and Welch 1996), in which the key assumption is that it is more difficult to understand signals from a culturally dissimilar group (Lang 1986). Also, researchers (e.g., Greenwald, McGhee, and Schwartz 1998, Bertrand, Chugh, and Mullainathan 2005, Stanley, Phelps, and Banaji 2008) have argued that due to negative unconscious attitudes—"implicit discrimination"—people often use simple decision rules biased against negatively stereotyped groups, which may result in little effortful scrutiny of relevant information. In our model, differences in acquired knowledge are an outcome of the agent's choice and can arise even if the signals provided are equally informative across groups and there are no unconscious biases in attention. This approach relates our model to a growing literature on rational inattention that uses an optimizing framework to study the effects of limited attention to the available information on a range of

meet but also analyze the speed of their reply. Conditional on receiving a callback, in our experiments we do not find any significant difference in response speed across ethnic groups.

⁹The effort to better inform theories of discrimination by collecting novel types of data and performing experiments across distinct markets relates our work to List (2004), who combines a natural field experiment with artefactual field experiments to distinguish between taste-based and statistical discrimination in a product market, and to Gneezy, List, and Price (2012), who measure discrimination based on disability, gender, race, and sexual orientation across several markets to understand how the controllability of a group attribute affects discrimination.

¹⁰Since the invitation decision combines a choice to learn more about an applicant with a pre-selection decision (narrowing down the pool of applicants), it is difficult to infer the sign of the gap in willingness to acquire information from observed gaps in the likelihood of invitation. This is particularly the case in lemon-dropping markets, where unfavorable stereotypes or preferences are predicted to lead to greater information acquisition but a lower likelihood of invitation.

(mostly macroeconomic) phenomena (e.g., Sims 2003, Mackowiak and Wiederholt 2009, Woodford 2009, Nieuwerburgh and Veldkamp 2010, Matějka and Sims 2011, Caplin and Dean 2015, Matějka and McKay 2015).

The rest of the paper is organized as follows. In Section 3.2 we develop a model of an agent who decides how much to learn about an applicant and we describe how "attention discrimination" can arise and its implications for discrimination in selection decisions. We also formulate testable predictions for the field experiments. Sections 3.3-3.5 detail the experimental designs and present empirical results in the rental housing and labor markets. Section 3.6 provides a discussion about how the results map on the proposed model and alternative interpretations. Section 3.7 concludes.

3.2 The Model of Attention Discrimination

3.2.1 Set-up of the Model

We model a two-stage decision-maker's (DM) choice about an applicant. A notable difference from existing models is that the level of additional information on the individual-specific quality is endogenous to the group's characteristics. In the first stage, the DM first observes the applicant's group of ethnic origin G , and then decides whether to pay additional attention to the applicant and whether to invite the applicant for an interview. In the case that the applicant is invited to the second stage, then the DM receives additional information about the applicant and chooses whether to accept him or not. The role of the first stage is to pre-select applicants to potentially save on costs from inviting unsuitable applicants to the second stage.

For the DM, the applicant is of an inherent unknown payoff π , which consists of two components:

$$\pi = q - d_G$$

where q is an unknown objective quality of the applicant which can include skill, work ethic or reliability, and d_G is the DM's known distaste towards the applicant's group G or the distaste of individuals with whom the DM interacts, e.g. customers or neighbors. Quality in group G is distributed according to $N(q_G, \sigma_G^2)$, which is known by the DM and it forms the DM's prior knowledge about q . With respect to information acquisition, the quality q can be expressed as follows:

$$q = q_G + q_1 + q_2,$$

where $q_1 + q_2$ is the deviation of the applicant's quality from the group's mean q_G . We assume that at the beginning of the whole process q_G is observed. Then, in the first stage the DM can acquire information about q_1 only, which is drawn from $N(q_1, \sigma_{G,1}^2)$ and is independent from q_2 . For instance, in the case of a job application, q_1 summarizes all quality that can be inferred from a resume. In the second stage (e.g. during the interview), the quality q is observable.

The DM knows what is the best alternative to the applicant, and thus knows the reservation payoff R from rejecting the applicant in either of the stages.¹¹ The DM maximizes the expectation of the payoff from accepting or rejecting the applicant less the incurred costs of inspection during the whole process.

$$payoff = \begin{cases} \pi & \text{if the DM accepts the applicant} \\ R & \text{if the DM rejects the applicant} \end{cases} - \text{inspection costs.}$$

In the first stage, the DM faces two choices. First, he chooses whether to pay the cost of inspection C_1 , e.g. whether to read the applicant's resume. If he pays the cost then he observes q_1 , and his posterior belief about the quality is $N(q_G + q_1, \sigma_G^2 - \sigma_{G,1}^2)$, since what is left to learn is q_2 only, otherwise the belief is $N(q_G, \sigma_G^2)$. Second, based upon the posterior belief he chooses whether to invite the applicant to the second stage, or not. The cost of invitation is C_2 . In the second stage, the DM observes the applicant's quality and decides whether to accept him. At this stage, when all costs of information acquisition are sunk, the applicant is accepted if and only if $q - d_G > R$.

DEFINITION (the DM's first-stage problem) *Upon observing G , the DM first chooses whether to incur C_1 and receive additional information, or to reject or invite the applicant without it. He chooses the action that maximizes the expected payoff.*

$$payoff(reject) = R$$

$$payoff(invite) = E[\max(R, q - d_G)] - C_2$$

¹¹The quality q and reservation payoff R also summarize all payoff-relevant implications given by the current market situation, which include the general equilibrium effects or even wage demands by each particular applicant. For instance, if in equilibrium everyone pays more attention to the majority, and filter out good majority candidates while the good minority applicants are still available, then q and R adjust accordingly.

$$payoff(info) = E[\max(R, E[\max(R, q - d_G)|q_1] - C_2)] - C_1$$

In principle, there can be two types of situations. A cherry-picking market is a selective one in which, without any information except for the group attribute, the DM prefers rejecting the applicant to inviting him to the second stage, $payoff(reject) > payoff(invite)$, and vice versa for the lemon-dropping market. For instance, a cherry-picking market is a labor market with many applicants for one job posting, where a priori a very few applicants are fit for the job, while in many locations the rental housing market is a lemon-dropping market, where an average applicant is acceptable.

Let us emphasize that the qualitative implications of the model would be unchanged if the DM faced a random pool of alternatives to the applicant in the second stage, rather than a given R . This is because the effect of a random pool of alternative applicants is already encompassed in the random component q_2 , which is observed during the second stage. The only thing that matters in the second stage is payoff from the applicant relative to payoff of the alternative; a high draw of q_2 can model a low draw of quality of the alternatives. The model thus also describes the behavior of a DM facing a sequential choice of applicants for a limited number of interview slots. In that case, R and the distribution of q_2 would depend on applicants the DM interacted with previously. For instance, if one applicant is accepted, then the next applicant faces a higher R .

3.2.2 Effects of Preferences and Beliefs on Attention Allocation

In this sub-section we describe how attention to an applicant is affected by sources of discrimination highlighted by prominent theories of discrimination. In the next sub-section we describe how endogenous attention affects the magnitude of discrimination in final selection decisions.

Proposition 1 below describes a new channel through which discrimination can operate: costly attention. It addresses how endogenous attention depends on the DM's choice situation and beliefs.¹² We discuss these predictions below, and test them in the empirical part of the paper.

¹²Note that the model describes the choice between no attention and some attention, only, and not between different positive levels of attention. We do this for the sake of simplicity as well as because most of our empirical results correspond to such a choice. In an alternative model with a sequential choice of levels of attention, where after some information is acquired the DM could choose to acquire more information, the results of Proposition 1 would hold, too. The only difference would be that the type of market would be conditioned on the information received before the choice of additional information.

PROPOSITION 1 (attention discrimination)

- A *Applicants from group G that are less attractive a priori (due to lower q_G , σ_G^2 , or higher d_G, C_2) are paid (weakly) less attention in the cherry-picking markets and (weakly) more attention in the lemon-dropping markets.*
- B *Applicants from a dissimilar group G with higher cost of attention C_1 or lower $\sigma_{G,1}^2$ are paid less attention in either market.*

Proof: Supplementary material.

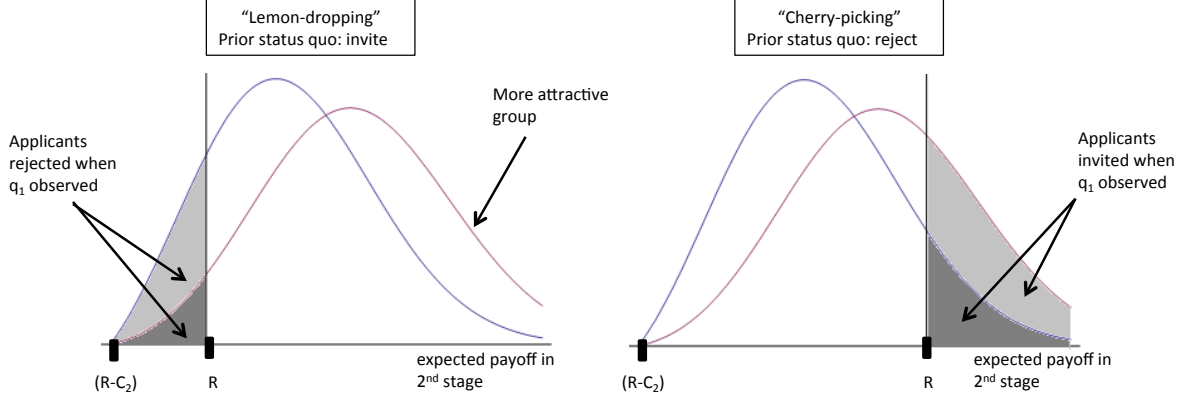
A distaste towards a certain group in our model is captured by parameter d_G (Becker 1971).¹³ An increase in d_G decreases the mean payoff from an applicant from group G . The resulting effect on attention differs across markets: a higher distaste implies less attention in the first stage in cherry-picking markets, and more attention in lemon-dropping markets. The reason is that costly information is useful only when there is a chance that it changes the status-quo decision (which is to reject in cherry-picking and to invite in lemon-dropping markets). Benefits from information are illustrated by the shaded regions in Figure 3.1. The figure presents distributions of expected payoffs from inviting the applicant to the second stage. These distributions are determined by distributions of q_1 and by expected payoffs in the second stage for each particular q_1 . In cherry-picking markets, information is useful only when the DM identifies applicants that are better not rejected. When the distaste d_G increases, then the distribution shifts to the left, since there are fewer good candidates, and the benefits from information decrease (left part). In contrast, in lemon-dropping markets (right part), benefits from information acquisition are given by the potential of discovery of bad applicants and thus a decrease in the mean payoff increases the DM's attention.¹⁴

Next, we consider statistical discrimination (Phelps 1972; Arrow 1973), which is driven by differences in beliefs about the applicant's quality. In our model, this channel is represented by a change in the mean q_G or variance σ_G^2 . The implications of a reduction in q_G are the same as for an increase in the distaste parameter d_G : less attention in cherry-picking and more attention in lemon-dropping markets. A decrease in the variance

¹³An alternative way of introducing differences in taste would be to vary reservation payoff R across groups. Note that an increase in R has the same effect as an increase in d_G of the same size.

¹⁴An increase in parameter C_2 has similar effects as an increase in d_G . This makes sense intuitively—higher C_2 may arise due to distaste towards interacting with a certain group during the interview. In our model, C_2 is not a cost of additional information that can be incurred independently of a selection decision, and unlike C_1 it is a cost that needs to be incurred for any accepted applicant and thus deducted from the payoff.

Figure 3.1: Expected Benefits from Information Acquisition in the First Stage



of beliefs σ_G^2 , when holding $\sigma_{G,1}^2$ fixed, has the same effects as a decrease in q_G , since a higher variance increases the likelihood of good candidates, while the bad ones are filtered out by the DM in the second stage.¹⁵

Last, we consider the effects on attention of a greater difficulty to understand signals from a culturally dissimilar group (Cornell and Welch 1996). In our model, such a dissimilar group would in the first stage be characterized by either a lower $\sigma_{G,1}^2$ (i.e., less uncertainty can be resolved by reading a resume) or by a higher cost of information C_{g1} (i.e., reading requires more effort). In both cases, attention in the first stage weakly decreases in either market. This is because both of these characteristics affect $\text{payoff}(\text{info})$ only: they determine the level of attention in the first stage, and do not affect the DM's choices when no information is provided in the first stage.¹⁶

¹⁵If σ_G^2 increases, then $E[\max(R, q - d_G)|q_1] - C_2$ increases for all q_1 , since the bad candidates are filtered out while the good ones are accepted, and thus the distribution in Figure 3.1 shifts to the right in the sense of first-order stochastic dominance.

¹⁶In contrast, characteristics described above ($d_G, q_G, \sigma_G^2, C_2$) affect not only $\text{payoff}(\text{info})$, but also $\text{payoff}(\text{invite})$, and thus determine the a priori attractiveness of the group.

To summarize, there are two types of group characteristics entering the model. Changes in characteristics related to taste and beliefs about groups ($d_G, q_G, \sigma_G^2, C_2$) affect a priori attractiveness of an applicant and have opposite effects on attention across the two types of markets. On the other hand, changes in characteristics related to a process of screening ($C_1, \sigma_{G,1}^2$) have the same effects in both markets.

3.2.3 Endogenous Attention and Discrimination in Selection Decisions

In this sub-section, we describe how endogenous attention can exacerbate discrimination in selection decisions, discuss cases when this is less likely to happen, and point to potential policy implications.

COROLLARY 1 (discrimination exacerbation)

- A *If both groups are either in the cherry-picking market or both are in the lemon-dropping market, then endogenous attention further disadvantages the group less attractive a priori (due to lower $q_G, \sigma_{G,2}^2$, or higher C_2) in the selection decision.*
- B *Endogenous attention disadvantages a dissimilar group G (with higher C_1 or lower $\sigma_{G,1}^2$) in the cherry-picking market and helps it the lemon-dropping market.*

Proof: Supplementary material.

The statement that endogenous attention disadvantages the less attractive group means that the difference in acceptance probability between applicants from a majority group and group G is (weakly) higher than if the level of attention to G were exogenously fixed at the level of attention paid to the majority.

The findings above imply that when the general population as well as group G face the same type of market (cherry-picking or lemon-dropping), then endogenous attention magnifies the effects of differences in taste and beliefs about groups. In cherry-picking markets, the applicant is rejected when no additional information about him is acquired and therefore the DM's attention weakly increases an applicant's chances of being invited. It follows that a higher chance of being invited implies a higher chance of being selected in the second stage, since the invitation is a prerequisite for selection and qualities are observed in the second stage. At the same time, the less attractive groups (i.e., groups with higher d_G , lower q_G , or lower σ_G^2) are paid less attention. In lemon-dropping markets,

attention decreases the likelihood of an invitation, and the disadvantaged group is paid more attention.

There is one special type of situation when endogenous attention may not magnify discrimination in selection decisions—a "middle-market", in which preferences for or beliefs about the two groups are very different and the less attractive group faces the selective situation and would be rejected in the absence of additional information, while the other group is in the lemon-dropping market. The disadvantaged group can, in this case, be paid more attention than the majority group, and endogenous attention can work to the minority's advantage.

The endogenous attention also magnifies the role of differences in difficulty to understand signals across groups in cherry-picking markets, where attention is desirable, but culturally dissimilar groups receive less of it. This is not the case in lemon-dropping markets, in which dissimilar groups also receive less attention, but here less attention improves chances of selection.

In the model, for the sake of simplicity, we assume that quality q is perfectly observable in the second stage. Thus, the DM's belief about an applicant's quality in the second stage is independent of the level of attention in the first stage, and hence endogenous attention in the first stage influences a final selection decision only via its effects on invitations to the second stage. If we allowed for the realistic case of imperfect knowledge in the second stage, then the magnifying effect of endogenous attention on discrimination in selection decisions would be further reinforced in cherry-picking markets. However, in lemon-dropping markets, while endogenous attention disadvantages applicants from the group G in terms of invitations to the second stage, the higher attention in the first stage may provide them with an advantage conditional on being invited. The DM would in the second stage possess more precise knowledge about such candidates, which might increase the likelihood of selection.

The findings above suggest the important role of timing of when the group attribute is revealed during the decision-making process in selective markets, an insight that is potentially interesting for policy. The following corollary is an immediate implication of Proposition 1 and Corollary 1.

COROLLARY 2 (timing of ethnic group revelation) *If both groups are either in the cherry-picking market or both are in the lemon-dropping market, then the probability that an applicant from a less attractive group is accepted is (weakly) lower if he is known to*

be from G a priori rather than when he is first considered to be from a general population and his membership in G is revealed only before the final selection decision.

Postponing the revelation helps the disadvantaged group by leveling the attention a DM pays to applicants. The probability that an applicant from a less attractive group is accepted is lower if he is known to be from group G prior to when the DM chooses whether to inspect the applicant in the first stage rather than when the applicant is first considered to be from a general population, and membership in G is revealed only before the final selection decision. This effect is not present in the standard model of statistical discrimination, because there the DM receives signals of exogenously-given precision and forms his posterior knowledge independent of the signals' succession, while in our model the first signal—i.e., the group attribute—affects the choice of whether to acquire an additional signal.

An important question that goes beyond the presented model is what are the dynamic and general equilibrium effects of endogenous attention? Endogenous attention has interesting implications for the persistence of discrimination that is driven by different beliefs across groups. It is known that such discrimination can persist in the long run, for instance if agents can invest in their skills, and if the skills and the investment are not perfectly observed (e.g., Coate and Loury 1993). In this case, a negatively-stereotyped group has less incentive to acquire the skills, which results in a self-fulfilling negative stereotype. With endogenous attention, this disincentive effect is further re-enforced in cherry-picking markets, where negatively stereotyped groups face not just lower likelihoods of acceptance but they are also less rewarded for their credentials due to lower attention given to them. The effect is attenuated on the lemon-dropping market. Regarding taste-based discrimination, endogenous attention does not seem to provide novel implications in terms of whether market forces would attenuate or eliminate such discrimination either by growth or endogenous entry of non-discriminating firms (Becker 1971; Arrow 1973).

Finally, the model also speaks to a dynamic setting, in which the DM is aware of the possibility of having inaccurate beliefs.¹⁷ In our model, the DM pays more attention when the uncertainty $\sigma_{G,1}^2$, which can be resolved in the first stage, is higher. This

¹⁷Alternatively, the DM can have inaccurate beliefs and be unaware of it. This may arise, for example, when DMs recall only a group's most representative or distinctive types (Bordalo et al.). Importantly, the effects on attention are driven purely by the form of beliefs, regardless of whether they are accurate or not. In selection decisions true qualities interact with beliefs and therefore inaccurate beliefs may change the predicted effects in either direction. For instance, being deemed as a highly homogeneous group is favorable in lemon-dropping markets, but disadvantageous in cherry-picking markets.

intuition extends to a dynamic model, too, when the whole selection process regarding one candidate plays the role of a first stage for all future candidates. When uncertainty about a group is higher, then the DM would pay more attention and be more likely to invite an applicant from group G simply to learn more about the whole group, and use such information in the future, i.e., to make future beliefs more accurate.

3.3 Field Experiment in the Rental Housing Market

In the first experiment, we study ethnic discrimination in the rental housing market in the Czech Republic, a market with a low level of selectivity (as we document below, a large fraction of applicants are invited).

We focus on two ethnic minorities: Roma and Asian. The Roma population constitutes the largest ethnic minority in the European Union (estimated at 6 million people, 1.2 percent) as well as in the Czech Republic (1.5-3 percent). Intolerance and social exclusion of Roma is considered one of the most pressing social and human rights issues in the European Union (Commission 2010). East Asians (mostly Vietnamese but also Chinese or Japanese) are the second-largest ethnic minority group in the Czech Republic (0.6 percent) and migrants from East Asia form large minority groups in many European countries. In the Czech Republic they are mostly self-employed in trade and sales businesses and lack formal employment.

Both minority groups are disadvantaged economically and socially, and face unfavorable stereotypes. The unemployment rate of Roma in the Czech Republic was estimated at 38 percent, compared to 9.4 percent overall unemployment rate in 2012. While 84 percent of the majority population complete a high school or university degree, the proportion is 47 percent and 33 percent for the Vietnamese and Roma adult population, respectively (Czech Statistical Office 2011). An opinion poll revealed that 86 percent and 61 percent of Czechs would not feel comfortable or would find it unacceptable to have Roma and Vietnamese as neighbors, respectively.¹⁸ In an online survey that is discussed below (Survey I), landlords expect individuals with Roma and Asian names to be worse tenants than the majority applicants for apartment rentals.¹⁹

¹⁸For more details about the socio-economic status of Roma in Central and Eastern European countries see Barany (2002). FRA & UNDP (2012) describe documented inequalities in education, employment, health and housing outcomes between Roma and majority populations in the Czech Republic and other EU countries. Spaan, Hillmann, and van Naerssen (2005) provide a detailed description of the integration of immigrants from East Asia in Europe.

¹⁹Similarly, another survey (Survey III) documents that university students expect individuals with

3.3.1 Experimental Design

Manipulating Identity of Applicant

The experiment was based on sending emails expressing interest in arranging an apartment viewing. To evoke ethnic minority status we designed three fictitious applicants: representatives of the Asian and Roma ethnic minorities and a control identity of the White majority group. The only real attributes of these identities were a name, an email address and a personal website.²⁰ We selected the names based on name frequency data: Jiří Hájek (White majority-sounding name), Phan Quyet Nguyen (Asian-sounding name) and Gejza Horváth (Roma-sounding name).²¹ Since the email address contained applicant's name, the name is arguably the first piece of information, which a landlord learns about the applicant. For the sake of brevity, we denote applicants with a White majority-sounding name as "White applicants" or as "majority applicants", applicants with ethnic minority-sounding names (both Asian and Roma) as "minority applicants", and applicants with Asian-sounding and Roma-sounding names as "Asian applicants" and "Roma applicants", respectively. Note that technically the results of our experiments describe the effects of the ethnic sounding-ness of the names rather than the effects of ethnicity itself.

To verify that landlords associated the selected names with respective ethnic groups, we conducted a pre-survey on a sample of 50 respondents. All respondents associated the name Jiří Hájek with the Czech nationality and the name Phan Quyet Nguyen with one of the Asian nationalities (92 percent associated it with the Vietnamese nationality), and the name Gejza Horváth was thought to be a Roma name in 82 percent of cases, indicating

Roma and Asian names to have a lower socio-economic status, as measured by education level and quality of housing.

²⁰There is a difficult trade-off involved in organizing this type of experiment. While informed consent is clearly desirable, it is extremely difficult to measure discrimination with the consent of participants in natural field experiments (List and Rasul 2011). Given their social benefits, audit studies and correspondence tests are considered among the prime candidates for the relaxation of informed consent (Riach and Rich 2002; Pager 2007). Our research has been approved by the Director of the Institute of Economic Studies, Charles University in Prague. We followed the conventional IRB standards for these types of experiments and took the following steps to minimize the landlords' costs and risks. In particular, the information acquisition was designed such that it took little effort and time and we quickly and politely declined invitations for an apartment viewing, within two days at most. We sent only one application to each landlord, and thus are not able to identify discrimination at an individual level, and after the data collection we deleted identifiers of individual landlords. A similar practice was followed in our companion experiments in the labor markets.

²¹Jiří is the most frequent Czech first name and Hájek is among the top 20 most frequent surnames in the Czech Republic. Nguyen and Horváth are the most frequent surnames for the Asian and Roma minorities, respectively.

a strong link between names and ethnicity. To confirm that the application emails from all applicants would be delivered and not identified as spam, prior to the implementation of the experiment we sent each variant of the email message to 40 individuals with email accounts from different providers. In all cases the emails were delivered successfully.

Manipulating Access to Information

In application emails, we used three manipulations of access to information about applicants (for an overview of the experimental design see Table 3.A.1 in the Supplementary material). First, in the No Information Treatment, the email contains a greeting and the applicant's interest in renting an apartment, but does not provide any information about the characteristics of the applicant other than his minority/majority-sounding name.²² Invitation rates in this treatment are informative about the type of market, since landlords can make inferences based on the applicant's name (and the short text) only. Recall that we defined the cherry-picking market as one where the status quo after learning a group attribute (but no other information) is not to invite an applicant—thus with no heterogeneity in the DM's thresholds the invitation rate should be close to zero. In contrast, in the lemon-dropping case the status quo is to select all applicants. We consider the invitation rate of 50 percent as an approximate dividing line between the two types of markets. This is the exact dividing line when heterogeneity among DM's is small and symmetric. In this situation a DM hesitates most whether to invite an applicant or not, and thus additional information is predicted to be most beneficial.

Second, in the Monitored Information Treatment, the email uses the same sentence to express interest in viewing an apartment as in the No Information Treatment. The only difference is that it includes the hyperlink to a personal website located in the applicant's electronic signature, which gives landlords an opportunity to acquire more information about an applicant. The link has a hidden unique ID number assigned to each landlord, which allows us to distinguish landlords who decide to acquire information about the applicant. Software similar to Mouselab monitors landlords' information acquisition on the website. Five different boxes are located in the main section of the website, each with a heading representing a type of information that is hidden "behind" the box—age, marital status, smoking habits, occupation, and education. A snapshot is displayed in

²²Specifically, the text of the email in the No Information Treatment was as follows: "Dear Sir/Madam, I am writing because I am very interested in renting the apartment that you have advertised. When would be a good time to come see the apartment? Best regards, Phan Quyet Nguyen". We avoided syntax or spelling mistakes. For wording of all manipulations of the email, see the Supplementary material.

Figure 3.A.1 in the Supplementary material. When the boxes were uncovered, landlords learned that the applicant is 30 years old, single, a non-smoker, and working in trade with a steady income. We randomly varied whether an applicant reported having a high-school or college degree.

Since only one box can be opened by a computer mouse at one point in time, the software allows us to identify whether a landlord decides to acquire information on an applicant's website, and how many and which pieces of information receive attention. These monitoring features provide direct insight into the process of information acquisition. In addition to the boxes with personal information, the website also contains tags for a personal blog, pictures and contact information (when accessed, an "under construction" note pops up, to reduce landlord's costs by limiting the time spent on the website). The design of the website is based on a professionally created template, which is freely available on the Internet.²³

In the third manipulation, the email again uses the same introductory sentence as in the No Information Treatment, but instead of providing a hyperlink to a website, the applicant reports the same characteristics directly in the body of the email. Specifically, we added the following text: "I am a thirty-year-old man, I am single, I have a college [a high-school] degree, and I do not smoke. I have a steady job (with a regular paycheck) at a company." Again, we randomly varied the education level. This allows us to study the effects of name on how much landlords respond to changes in available information in terms of invitation rates. The motivation for this treatment is to test whether name effects on responsiveness mimic the name effects on attention from the Monitored Information Treatment.

An online survey (Survey I) implemented after the experiment among a different sample of landlords (N=60) shows that the set of applicant characteristics reported in the second and third manipulations are considered attractive, as compared to the typical population of applicants on this market. The landlords (N=60) were given two profiles (across subjects), which contained the same set of characteristics as described above, and asked: "Based on your experience with renting an apartment, how would you compare the following applicant to other applicants? 1=strongly above average, 2=above average, 3=average, 4=below average, 5=strongly below average". We find that both profiles,

²³Still, to some landlords the website may appear unusual and this may affect their callback. Nevertheless, it should be noted that the content and the design of the website cannot affect a decision as to whether or not to open it, since the decision happens when the landlord sees only the link. Providing a hyperlink to a personal website is a common feature in an electronic signature.

those with a high-school and a college degree, were evaluated as substantially above average (2.19 and 1.64, respectively).²⁴

3.3.2 Sample Selection and Data

The experiment was implemented between December 2009 and August 2010 in the Czech Republic, mostly in Prague. Over that period, we monitored four (out of ten) major websites that provide rental advertisements. Placing an ad on these websites requires a small fee, while responding to an advertisement is free. We chose to apply only for small homogenous apartments of up to two rooms with a separate kitchen that look suitable for a single tenant without a family. We excluded offers mediated by real estate agents and also offers where landlords did not make their email publicly available and relied on a telephone or an online form (66 percent), in order to be able to monitor information acquisition. Overall, we responded to 1800 rental ads and randomly assigned an applicant name and provided information. We recorded the gender of the landlord, implied by the name, and the characteristics of apartments commonly published as a part of the advertisement such as rental price, the size of the apartment and whether it is furnished. These characteristics vary little across experimental treatments, indicating that randomization was successful (Table 3.A.3).

To measure attention in the Monitored Information Treatment, we record whether a landlord visits an applicant's personal website and how many and which boxes with information he uncovers. To measure responses to the applicant, we distinguish between a positive response, indicating either a direct invitation to an apartment viewing or an interest in further contact, and a negative response, capturing the rejection of an applicant or the absence of response.²⁵ Note that with the correspondence experimental approach a researcher does not measure the ultimate outcomes, i.e. whether an applicant rents the apartment and for what price. Nevertheless, since the invitation is typically a prerequisite for the final positive decision, it is likely that the gaps in the share of positive responses across ethnic groups translate into gaps in final decisions about actual rental.

²⁴These characteristics are likely seen particularly positively (relative to expectations) for minority applicants, given the gaps in education and employment relative to the majority population.

²⁵As a robustness check, we also estimated the effect of minority-signaling names on callback (Table 3.A.4), which distinguishes applications that result in contact, regardless of whether it is a positive or negative response. Overall, we find a qualitatively similar impact of names on the callback rate as on the invitation rate.

3.3.3 Results

Do Landlords Discriminate Against Minorities?

We start the analysis by looking at whether ethnic minorities are discriminated against when no information about the applicant other than his name is available to a landlord (No Information Treatment). In this treatment, the invitation rates reflect the tastes and prior beliefs about the expected characteristics of each group. We find that majority applicants are invited for an apartment viewing in 78 percent of cases, while minority applicants receive invitations in only 41 percent of cases (Panel A of Table 3.1). The gap that arises solely due to name manipulation is large in magnitude (37 percentage points, or 90 percent) and statistically significant at the 1 percent level. Put differently, minority applicants have to respond to almost twice as many advertisements to receive the same number of invitations as majority applicants.

Next, we distinguish between applicants with Asian- and Roma-sounding names. The invitation rates are very similar: 43 percent for the Roma minority and 39 percent for the Asian minority applicants. The difference in invitation rate between the two minority groups is not statistically distinguishable (Column 8), while the gap between the majority and each of the two minority groups is large and similar in magnitude (Columns 5 and 7). Table 3.2 documents the findings in a regression framework, where we control for the landlord's gender and the characteristics of the apartment described in an advertisement (price, size, furnishings).

Observation 1: Applicants with minority-sounding names are discriminated against. If no information about applicants is available, applicants with a majority-sounding name are 90 percent more likely to be invited for an apartment viewing compared to applicants with a minority-sounding name.

In the model, information acquisition is the most valuable and the DM pays the most attention when ex ante expected payoffs from invitation and rejection are equal (the invitation rate without additional information is 50 percent), and thus when the DM hesitates ex ante then any piece of information is useful and can affect the decision. The further the invitation rate is from 50 percent the less information is acquired. The invitation rate for the majority applicants (78 percent) suggests they are in the lemon-dropping situation and the mean of the prior belief about this group is far above the threshold level of quality necessary for invitation. On the other hand, the invitation

rate for minority applicants is 41 percent. Given that some fraction of apartments might have already been rented out by the time we sent the email and thus their owners were unlikely to invite any applicant, it is difficult to say for sure whether the prior mean about a minority group is above (lemon-dropping situation) or below (cherry-picking situation) the threshold. Importantly, though, it's clear the landlords hesitate more about whether to invite applicants from the minority group and thus acquiring more information about minority applicants should be more valuable, compared to acquiring information about majority applicants.

Do Landlords Choose Different Levels of Attention to Information Based on the Ethnicity of an Applicant?

In the Monitored Information Treatment, we find that only less than half of the landlords open the applicant's website even though the cost of acquiring information is very small—literally one click on the hyperlink. Importantly, the applicant's name matters for attention allocation (Panel B of Table 3.1 and Table 3.3). While 41 percent of landlords opened the website of minority applicants, 33 percent did so for majority applicants. When summing the number of applicant characteristics to which a landlord pays attention (the maximum is five), we find that landlords learn about 1.75 characteristics of a minority applicant and 1.29 for a majority applicant. Similarly, the likelihood of opening at least one of the boxes with information is 40 percent for minority and 30 percent for majority applicants, and the likelihood of opening all the boxes is 26 percent for minority and 19 percent for majority applicants. These differences in information acquisition measures across an applicant's ethnicity are statistically significant and are driven by both a greater likelihood of opening the website as well as more effort to acquire information, conditional on opening the website. Among a sub-sample of landlords who opened an applicant's website, we still observe that landlords are significantly more likely to open at least one of the boxes with information and to open a higher number of boxes when the applicants have minority names compared to majority name.

Table 3.A.5 reports further results about how ethnicity affects the process of information gathering. While the name affects the amount of information acquired, we do not find a systematic influence on which type of information is acquired as well as on the order in which different pieces of information are acquired. Unconditional on opening the website, the likelihood of opening a box about, for example, education level is 36

percent for minority applicants and 27 percent for majority applicants (Panel A). Thus, the difference due to name manipulation is 33 percent (or 9 percentage points). A similar picture arises for other individual characteristics: the likelihood of paying attention to those is 30-46 percent greater for minority applicants compared to majority applicants. Also, the landlords who visit the website are more likely to open each of the boxes for minority applicants compared to majority applicants, but the differences are not statistically significant with the exception of the box with occupation information (Panel B).

In terms of the order of uncovering the boxes, we find that conditional on opening the website, the likelihood of uncovering each of the boxes as the first one does not significantly differ across ethnicity (Panel C). Similarly, conditional on opening all the boxes, the order of uncovering does not differ across applicants' names (Panel D). Together, these results suggest that the observed differences in acquired information are not driven by landlords being worried about a particular single attribute of minority applicants, but rather by a more general effort to screen this group more carefully.

Distinguishing between the two minority groups reveals that, compared to the majority applicant, landlords acquire more information about both Roma and Asian applicants (Columns 5 and 7 of Table 1). We also observe that the amount of acquired information is somewhat (although insignificantly) greater for Roma applicants relative to Asian applicants (Column 8). This is interesting given that the landlords appeared to hesitate most on whether to invite Roma applicants, since the invitation rate of this minority was closest to the 50 percent invitation rate. Table 3.3 documents the findings in a regression framework.

Observation 2: Landlords pay more attention to available information about applicants with a minority-sounding name relative to applicants with a majority-sounding name.

Responsiveness to Available Information

In order to test whether landlords are more responsive to available information provided by minority applicants compared to majority applicants, as suggested by observed differences in attention, we estimate the effects of three manipulations in the available information on invitation rate: (1) adding a sentence to the email message signaling attractive

characteristics of the applicant, (2) varying the education level between high-school degree and college degree in the added sentence, and (3) having access to an applicant's personal website.

We find that the invitation rate responds to information provided by applicants with minority names²⁶, a pattern which is consistent for all three manipulations of available information. Column 5 in Panel A of Table 3.2 shows that, relative to the No Information Treatment, the invitation rate increases by 8 percentage points for minority applicants who add the sentence reporting a high-school education. The increase is 15 percentage points for minority applicants who add a sentence and report having a college degree. The pure effect of reporting a college degree compared to a high-school degree is 8 percentage points, which is marginally significant statistically (Column 3 of Table 3.A.7). In contrast, there is little response in the invitation rate when the same manipulations of available information are performed by the applicant with the majority name. The invitation rate remains at the same level, 78 percent, independent of whether the applicant provides no information, includes a sentence about his characteristics, and also does not respond to changes in his education level (Column 4 of Table 3.2).

Taken together, the decision-makers are found to be more sensitive to information provided by minority candidates compared to majority candidates (Columns 3 of Table 3.2). The interaction effect of having a minority name and adding a sentence with a college degree is 14 percentage points. The interaction effect is still positive (8 percentage points) when reporting a high-school degree. Similarly, among applicants who provide information in the body of the email, the interaction effect of a minority name and reporting a college degree is 7 percentage points (Table 3.A.7). The last two interaction effects are not statistically significant. As a consequence, the discrimination in terms of invitation rate is 37 percentage points in the No Information Treatment, and it diminishes to 29 p.p. for applicants who add a sentence and report a high-school degree and to 22 p.p. for applicants who add a sentence and report a college degree (Table 3.A.7).

Observation 3: The landlords' decision whether to invite an applicant is responsive to manipulations of the available information about applicants with a minority-sounding name, while the decision is not (or only a little) affected by the same changes in the available information about applicants with a majority-sounding name.

²⁶Landlords are responsive to changes in available information about both minority groups (Table 3.A.6).

Giving access to a personal website leads to an increase of 8 percentage points in the invitation rate for minority applicants, while it causes a moderate but not statistically significant decrease of 6 percentage points for majority applicants. As a result, the gap in the invitation rate between majority and minority applicants decreases from 37 percentage points in the No Information Treatment to 23 percentage points in the Monitored Information Treatment (Column 3 of Table 1) and the difference is statistically significant (Column 3, Panel A of Table 3.2).

Next, we compare discrimination among landlords who choose to acquire information with those who don't in the Monitored Information Treatment. Note that since attention is not experimentally manipulated, the difference in decisions between these two groups cannot be interpreted causally because we cannot separate the effect of having more information from the self-selection of certain type of landlords. We find a positive relationship between opening a website and the likelihood of an invitation (Column 2, Panel B of Table 3.2) and this relationship is slightly higher for the minority candidates (Column 3).²⁷ The landlords who did open an applicant's website discriminate less than those who did not—the gaps in the likelihood of invitation are 18 and 29 percentage points, respectively—although the difference is not statistically significant (Column 3-5).

Last, among a group of landlords who chose to uncover box about education, reporting a college degree increases the invitation rate compared to reporting a high-school degree. Interestingly, conditional on paying attention, the effect is similar for minority and majority applicants (Columns 6-8 of Table 3.A.7).²⁸

Other Results

In order to gain some insight about landlords' priors and to better understand possible motivations for observed differences in attention, we conducted an online survey (Survey I) among 60 landlords. The landlords are drawn from the population of landlords who post rental offers online, but are different from those in our experiment.²⁹ Conditional only on name, we directly elicited the mean expected satisfaction with an applicant (in our model,

²⁷Notice that from the sign of this coefficient we cannot draw inferences about whether attention helps or hurts candidates on this market in general and compare that with the model's predictions, since here we measure a relationship between attention and selection for a candidate with specific (attractive) attributes and not for the whole population of candidates from a given group.

²⁸We don't find any systematic effect of education level on search patterns (Table 3.A.8).

²⁹We have asked 817 landlords to participate during the months of January and February 2015. The response rate was 7.3 percent. In total we have 89 observations, since some of the 60 landlords answered questions about two or three applicants with different names.

$q_G - d_G$), the variance of expected satisfaction (σ_G^2) and expected informativeness of an applicants' personal website ($\sigma_{G,1}^2, C_1$), since differences in each of these three parameters across groups are predicted to generate asymmetry in attention. Each landlord was given snapshots of a flat rental advertisement and of an email response used in the experiment. The first question was: "How likely it is that the following applicant would be a tenant with whom you would be: 1=highly dissatisfied, 2=somewhat dissatisfied, 3=neutral, 4=somewhat satisfied, 5=highly satisfied?" The landlords were asked to allocate ten tokens, each representing a ten percent probability, to the five options. This allows us to measure mean and variance of overall expected satisfaction at the individual level.³⁰ The second question was: "Imagine you have access to the personal website of the applicant. To what extent do you think the website is informative for evaluating him as a prospective tenant? 1=very uninformative (I will not learn much about an applicant from reading his website), 2=somewhat uninformative, 3=somewhat informative, 4=very informative (I will get a clear idea about the candidate from reading his website)."

Panel A of Table 3.A.9 shows the results. When compared to the majority name, both Asian and Roma names significantly reduce the mean of expected satisfaction. In contrast, we find virtually no effect of names on the standard deviation of expected satisfaction as well as the expected informativeness of a personal website. Taken at face value, these results support the interpretation that differences in information acquisition across groups observed in experiments are due to unfavorable preferences or prior means, but not due to greater uncertainty about minority candidates or the expected greater informativeness of their website. Nevertheless, these supporting findings need to be taken cautiously, since such direct questions are more vulnerable to social desirability bias and were answered by a sample of landlords different from the decision-makers in the experiment.

In principle, the observed greater inspection of personal websites of minority applicants in the housing market could be due to confirmation bias or due to pure curiosity to read about dissimilar individuals,³¹ both of which would imply that acquired information should affect selection decisions less for minority candidates. Additional results do not provide support for this interpretation: The correlation between opening a website and inviting is similar or greater for minority applicants compared to majority applicants,

³⁰We deliberately focus on measuring priors about overall satisfaction instead of priors about specific attributes, since previous work suggests that stereotypes about a group may vary across different dimensions (Fiske et al. 2002).

³¹Of the landlords in our sample 93 percent have a White majority-sounding name.

and the observed differences in attention across groups mimic observed differences in responsiveness to manipulations of available information.

To summarize the main results in the rental housing market, we find that negatively stereotyped names affect both the choices of whether to invite an applicant for an apartment viewing as well as the attention paid to information prior to this decision. Applicants with minority-sounding names are more thoroughly inspected and less likely to be invited for an apartment viewing. Differences in the observed level of inspection across the groups mimic greater responsiveness of the invitation rate to the manipulation of available information about the quality of applicants with minority-sounding names.

3.4 Field Experiment in the Labor Market—Czech Republic

The second experiment shifts the exploration of discrimination to the labor market. Here, we aim to study discrimination during a selection process in which decision makers pick only a few winners out of a large pool of applications, in contrast to the rental housing market.

3.4.1 Experimental Design

We use the same names as in the rental housing market experiment to evoke Asian, Roma and White majority ethnic status. The experiment was implemented between August and October 2012 in the Czech Republic. Over that period we monitored the major online job site (www.jobs.cz) and responded to online job advertisements. We implement the treatment with the monitoring of information acquisition, and send an application via email. The email contains a greeting, the applicant's interest in the job opening, his name and a hyperlink to his professional resume on a website.³²

We created a conventional resume, following real-life resumes, and we responded to job offers for which the applicant satisfied all the education and qualification criteria. The resume has six parts: education, experience, skills, hobbies, references, and contacts. Applicants are 30-year-old males, have prior work experience as administrative workers,

³²The exact wording was as follows: "Dear Sir/Madam, I am writing because I am very interested in job position advertised by your company. You can find my resume in this hyperlink: phanquyet-nguyen1982.sweb.cz. Best regards, Phan Quyet Nguyen".

and we randomly determined (orthogonally to name) whether they obtained a high-school or a college degree. They report a good knowledge of English, PC skills, and a driver's license. They also list their hobbies and provide two reference contacts. We believe the resume was roughly comparable to that of a standard applicant for these types of jobs.

When employers open the website, they can see a standard version of the resume. Further, they can click on "learn more" buttons placed below each resume category label (contact, education, experience, skills and hobbies). For example, when the website is accessed it reveals basic information about previous employment experience: the name of the firm, the position held and the time period. By clicking on the "learn more" button below the "Experience" label, the website reveals the applicant's responsibilities (document management, administrative support of consultants, work with PC). Thus, in addition to monitoring whether an employer opens the resume, we measure whether an employer decides to acquire more and which type of information. An example of the shorter as well as the expanded form of the resume is in Figure 3.A.2 in the Supplementary material.

3.4.2 Sample Selection and Data

We focused on job openings in sales, customer service, and administrative work. We selected these job categories because they have a sufficient flow of new openings and are similar enough not to require subtle adjustments of particular skills in resumes. In addition, we aimed to minimize the costs for the employers of reviewing the resumes and thus we selected job categories that involve less intensive inspection of applicants compared to higher-skill jobs. We also sent only one email to each employer and politely declined all invitations for job interviews within two days.

We target the population of employers who use the Internet to advertise job openings. To be able to monitor the opening of an applicant's resume, we had to exclude ads in which employers did not make their email publicly available and required applicants to call or use an online form (59 percent). Overall, we responded to 274 job openings and to each of them we randomly assigned the name of an applicant. We record the type of job, the job requirements and the time when the application was sent. The means of the observable characteristics of job openings are similar across the three groups of applicants (Table 3.A.10), with the exception of a somewhat higher likelihood of majority applicants applying for openings that required previous job experience, compared to minority appli-

cants (p-value=0.13) and the lower likelihood of majority applicants applying during the holiday period (p-value=0.12). In the analysis, we rely on a comparison of means across treatment conditions, as well as a regression analysis in which we control for observable characteristics.

We study how name manipulation affects two types of choices: attention to a resume and the selection decision. First, we measure whether an employer opened an applicant's resume by clicking on the hyperlink to a resume website. Further, we identify which additional information about an applicant an employer uncovered by clicking on the "learn more" buttons. As was the case in the first experiment, we do not measure the ultimate outcome of a selection process (an actual employment offer and wage). The outcome measure is whether the employer emailed or called the applicant back with a decision ("callback") and whether the employer decided to invite the applicant for an interview ("invitation"), a more precise outcome of the initial stage of selection process than callback.³³

3.4.3 Results

Are Ethnic Minorities Less Likely to be Invited for a Job Interview?

Panel A of Table 3.4 documents a large amount of discrimination against minority applicants. The callback rate for majority applicants is 43 percent and only 20 percent for minority applicants, making a difference of more than 100 percent, which is highly significant statistically ($p < 0.01$). A similar picture arises when we turn to the invitation rate. While majority applicants are invited in 14 percent of cases, minority applicants receive an invitation only in 6.3 percent of cases. The gap is statistically significant (p-value=0.03) and is large in magnitude (133 percent).

While we observe the almost identical treatment of applicants with Roma- and Asian-sounding names in the rental housing market, we find some differences in the labor market. Both minority groups are less likely to be invited for a job interview compared to the

³³Since the application was sent via email, the most common response from employers was also via email: 25.9 percent of employers emailed back, 9.1 percent invited the applicant for an interview and 16.8 percent declined the application. Employers could also call the applicant's cell phone number reported in the resume. However, only a few employers called back (5.8 percent). We recorded "missed calls" on each cell phone and then called back to determine the particular employer. Most of the employers who made a phone call also responded via email, and one employer sent a text message with an invitation; in only six cases we cannot directly identify whether the employer who called back meant to invite the applicant or not. In the main estimations we assume they did not, given the large fraction of declines in the email responses and the fact that these employers did not get in touch with the applicant via email.

majority group. The gap, however, is larger and more significant statistically for the Asian minority applicant (5.1 percent invitation rate, $p\text{-value}=0.03$) than the gap for the Roma minority applicant (7.8 percent invitation rate, $p\text{-value}=0.18$). Put differently, Asian applicants need to send 20 applications to receive one invitation, Roma applicants 12.5 and majority applicants 7.5. Columns 1-2 of Table 3.5 demonstrate the evidence in a regression framework. Controlling for observable job characteristics—required high school education, required previous experience, the type of job, and whether the application was sent during the summer holidays—does not affect the size of the observed gaps in the invitation rates and somewhat increases precision. We also find that employers who decide to read a resume are more likely to invite the applicant and those who request previous job experience are less likely to invite the applicant (available upon request).³⁴

Observation 4: Applicants with minority-sounding names are discriminated against in the labor market. An applicant with a majority-sounding name is 180 percent more likely to be invited for a job interview compared to an applicant with an Asian-sounding name, and 75 percent more likely compared to an applicant with a Roma-sounding name.

Additional results suggest that human resource managers do not reward applicants for having a higher level of education than requested in the advertisement. While the applicants reported to have either a high-school or a college degree, the positions they applied for requested a high school (80 percent) or lower education level. Specifically, conditional on opening the resume, the invitation rate does not differ for applicants who report having a high-school degree and for those who report a college degree, for both minority as well as majority applicants (Table 3.A.12).

The invitation rate (on average 9.1 percent) in the labor market is much lower than 50 percent, despite the fact that the resume signals the relatively high quality of applicants for the selected job types. If we were to link this observation to theory, it would imply that mean prior beliefs about the quality of all groups are below the threshold necessary for an interview invitation and that the labor market is the "cherry-picking" type of market.³⁵

³⁴It is noteworthy that discrimination is not restricted to jobs where language skills and interactions with customers are central (sales and services), and thus employers could presumably discriminate due to a belief about language use or due to the expected taste-based discrimination of their customers (Table 3.A.11).

³⁵Low invitation rates seem to be a ubiquitous feature of labor markets. The invitation rates do not get anywhere close to 50 percent in any segment of the market, which does not allow us to test the theoretical prediction about the "switch" in relative attention between minority and majority groups

Resumes of applicants with minority names, and with Asian names in particular, are thus predicted to receive less attention compared to resumes provided by majority applicants.

Do Employers Choose Different Levels of Attention to Information Based on the Ethnicity of an Applicant?

We start by looking at the likelihood of opening a resume. Of employers, 58 percent open the resume. Name again matters. We find that while 63.3 percent of employers visit the webpage with the resume of majority applicants, only 47.5 percent of employers do so when they receive an application with the Asian-minority name. The difference is large in magnitude (34 percent) and significant statistically (Panel B of Table 3.4 and Column 4 of Table 3.5) and it demonstrates that ethnicity signaled by name represents a barrier even at the very start of a selection process, before any information about an applicant is acquired. Moreover, in some firms it is common to delegate the printing of all received applications to an assistant, and printed resumes are then screened and evaluated by a different person. In such cases our experimental design fails to measure differences in attention, biasing down the estimated effect of a name on the likelihood of opening a resume. Regarding Roma-minority applicants, i.e. the group with the invitation rate between the majority and the Asian-minority applicants, we find no discrimination in attention: the likelihood of opening the resume is the same as for the majority applicant and higher compared to the Asian-minority applicant (Column 4 of Table 3.5).

Further, we study whether employers differentiate attention after opening the resume. Overall, we find relatively little interest to expand the resume; only 14 percent of employers in our sample clicked on at least one out of five "learn more" buttons and only 1 percent (3 employers) clicked on all buttons. Despite relatively little variation, the results (reported in Table 3.A.13) reveal that employers were somewhat less active in acquiring information about the Asian-minority applicant: they clicked on a lower number of "learn more" buttons and were less likely to click on all the buttons. This is mostly driven by a reduced interest in getting more detailed information about experience (13 percent for the majority and 4 percent for the Asian candidate) and skills (6 percent and 2 percent). Although these differences are large in magnitude, they are not significant statistically at conventional levels. Taken together, employers paid significantly less attention to the qualification of Asian-minority applicants: when considering majority applicants, 16 per-

within one type of market (in contrast to testing it across two different markets, as we do in this paper).

cent of employers further inspected at least one out of three categories that seem relevant for assessing qualifications (experience, education, and skills), while only 6 percent made that effort when considering Asian-minority applicants (Panel B of Table 3.4). We find no differences in acquiring information about contacts and hobbies.³⁶

In sum, the Asian minority, i.e. the group with the lowest invitation rate, receives the least attention, which is in line with the predictions of our model of attention discrimination. Since the invitation rate of the Roma-minority applicant is lower than the invitation rate of the majority applicant and higher than the invitation rate of the Asian-minority applicant, the model predicts that the amount of acquired information about the Roma applicant should also be somewhere between the Asian applicant and the White applicant. Nevertheless, we do not observe any differences in attention compared to the majority applicant.

Observation 5: Employers are 34 percent more likely to read a resume provided by applicants with majority-sounding names relative to applicants with Asian-sounding names. Conditional on opening a resume, employers more closely inspect the qualifications of applicants with a majority-sounding name relative to applicants with an Asian-sounding name. There is little difference in the likelihood of opening a resume as well as in the depth of resume inspection between applicants with majority- and Roma-sounding names.

Other Results

In order to get a better sense about priors of employers regarding candidates with the names used in our experiment, we conducted an online survey (Survey II) among 39 human resource managers.³⁷ Specifically, employers were shown a snapshot of a job advertisement (similar to one we responded to in the experiment) and of an email response, including the name of the applicant (but not resume). They were asked similar questions as in Survey I, which were adapted to the labor market context. To proxy mean and variance of expected satisfaction with an applicant, we asked: "How likely it is that the

³⁶The fact that only a few employers used the "learn more" buttons prevents us from analyzing differences in the order in which different types of information were acquired.

³⁷The sample of human resource managers in this survey is different from the sample in our correspondence test. The survey was implemented in February 2015. We have directly invited 913 human resource managers via email (response rate 4.3 percent). In total, we have 90 observations, since some of the 39 human resource managers answered questions about two or three applicants with different names.

following applicant would be an employee with whom you would be: 1=highly dissatisfied, 2=somewhat dissatisfied, 3=neutral, 4=somewhat satisfied, 5=highly satisfied?" Employers were asked to allocate ten tokens, each representing a ten percent probability, to the five options. The second question serves as a measure of the expected informativeness of applicant's resume: "Imagine you have access to a professional resume of the applicant. To what extent do you think the resume is informative for evaluating him as a prospective employee? 1=very uninformative (I will not learn much about an applicant from reading his resume), 2=somewhat uninformative, 3=somewhat informative, 4=very informative (I will get a clear idea about the candidate from reading his resume)."

Panel B of Table 3.A.9 shows the results. In terms of mean and variance of expected satisfaction, we find a similar pattern for employers as for landlords. On average, employers expect to be significantly less satisfied with minority compared to majority candidates, while the uncertainty about the candidates, in terms of variance, seems to be very similar across names. It is also noteworthy, in light of the lowest attention and invitation rate observed for the Asian applicant in the experiment, that employers have the worst expectations about this group, although the difference between the Asian and Roma applicants is not significant statistically. At the same time, we find that employers expect the resume of minority applicants to be less informative than that of the majority applicant and virtually no differences in this measure across the two minority groups. These results suggest that the observed discriminatory behavior in the experiment is unlikely to be motivated by a lower variance of beliefs about minorities ("you can't tell them apart"), but rather by negative stereotypes, preferences, or the lower expected informativeness of a resume.

Although the selected names used in the experiments strongly signal ethnicity, they may also signal some other characteristics, social background in particular. In order to assess whether our findings can be explained by discrimination against individuals with low socio-economic status (SES), we turn to the survey on perceptions, which we conducted among 92 university students (Survey III). For each name used in the experiment, we measure associations with level of schooling and quality of housing. We find that the majority-sounding name is perceived as having the highest SES, the Roma-sounding name the lowest SES, and that the Asian-sounding name is in the middle (Table 3.A.2). If the results in the experiment were fully driven by SES, we would expect the Roma minority to be the most discriminated against, and thus this analysis does not provide strong support for the SES interpretation of the name effects.

Motivated by the findings of previous research which documented differences in callbacks between names associated with the same ethnicity (Jacquemet and Yannelis 2012), in Survey III we also included three other names for each ethnicity besides the names used in the experiments. Within each ethnic group, all majority-sounding names and all Asian-sounding names are perceived very similarly (Table 3.A.14), suggesting our results are likely to be close to the average discrimination of people with Asian-sounding names. The Roma-sounding name used in our experiment is perceived similarly as one of the three names and as signaling a somewhat lower socio-economic status compared to the two remaining Roma names, indicating that the observed discrimination might be an overestimate of the average discrimination of the Roma minority.

3.5 Field Experiment in the Labor Market—Germany

The third field experiment tests two further considerations: (i) the generalizability of attention discrimination against negatively stereotyped ethnic groups to a different country, Germany, and (ii) generalizability to other types of signals of an applicant’s quality beyond ethnicity, such as being unemployed.

3.5.1 Experimental Design

We study discrimination against the Turkish minority in the German labor market. Individuals with a Turkish background represent the largest minority in the country (2.9 percent of population). Migrants from Turkey came to Germany predominantly in the 1960s and their children and grandchildren, raised in Germany, now represent a significant share of Germany’s labor force. The unemployment rate in 2012 among the majority German population was 6.2 percent, while it was 14.4 percent among immigrants. Importantly, Kaas and Manger (2012) found evidence of discrimination against the Turkish minority by employers. In their experiment a White majority-sounding name increased the likelihood of a callback by 14 percent compared to a Turkish-sounding name. We build on these results by focusing on the effects of a minority name on information acquisition prior to a selection decision to invite an applicant for a job interview or not.

We use the same names as Kaas and Manger (2012) to signal White majority and Turkish ethnicity.³⁸ The experiment was carried out between August and September 2013.

³⁸The first names and surnames of White majority applicants—Denis Langer and Tobias Hartmann—

We used major online job advertisement sites in Germany and sent email applications to 745 online job postings in sectors such as information and communication, administration, health and education, manufacturing and construction. We responded to all job ads that were posted directly by the company and included an email contact in the text of the posting (66 percent of all ads).

The email contains a greeting, the applicant's interest in the job opening, his name, and a hyperlink to his professional resume on a website. Orthogonally to name treatments, we have implemented three additional conditions by randomly varying the text of the application email. In the baseline condition (50 percent of responses) the text was the same as in the Czech Republic. Next, we have implemented two conditions (25 percent of responses each), in which the text of the application email contains a negative signal about an applicant's quality. Specifically, the email includes the following sentence: "I have been searching for a job for two months [a year and a half]." Both unemployment lengths are common in Germany: 45 percent of the unemployed in 2012 were unemployed for at least one year. The negative signals come from revealing the applicant's potential unemployment to the employer, as well as the carelessness demonstrated in releasing this information in the introductory email. Otherwise, the text is identical as in the baseline condition. Observable characteristics of the job openings vary little across experimental conditions (Table 3.A.15).

As outcomes of interest, we again focus on measures of information acquisition by the employers. We have altered the design of the experiment to aim at more detailed measures of effort to acquire a resume. The provision of a hyperlink allows us to distinguish whether an employer decides to open the resume. In this experiment clicking on the link does not, however, reveal an applicant's resume but instead a browser displays a message indicating a temporary error on the server.³⁹ We measure whether an employer clicked on the link, the number of times the employer attempted to re-open the resume and the likelihood that the employer sent an email requesting the resume to be re-sent, an action that requires nontrivial effort. Since the gap in terms of callback has already been established in previous work (Kaas and Manger 2012), we have not proceeded by responding back

belong to the 30 most common names in Germany. The names of the ethnic minority applicants—Fatih Yildiz and Serkan Sezer—are very common for male descendants of Turkish immigrants.

³⁹The message is "Database connection error (2): Could not connect to MySQL Server!" Such a message commonly appears when announcing a failure to connect to the database server due to technical problems on the provider's side and thus should not indicate a mistake on the part of the applicant. Nevertheless, we cannot rule out that the failure to open the resume was perceived by some employers as an applicant's mistake and thus a negative signal about his quality.

with a resume to minimize the costs on the part of employers, and we focus purely on information acquisition.

3.5.2 Results

We find that a minority name reduces employer’s effort to acquire information about an applicant compared to a majority applicant. This effect holds for all three measures of information acquisition: likelihood of opening an applicant’s resume (Table 3.6 and Column 1, Panel A of Table 3.7), number of attempts to open an applicant’s resume (Column 3) and a likelihood of writing back requesting an applicant to re-send the resume (Column 5).

Observation 6: Applications with Turkish minority names receive lower attention in all three measures than applications with majority names in the German labor market.

Interestingly, the gap is greater at higher levels of effort. The likelihood of clicking on the resume link at least once is 75 percent for minority applicants and it increases by 8 percent for majority applicants (to 81 percent). For the number of times an employer tried to open the resume the difference is 31 percent (2.1 for minority and 2.8 for majority applicants). Finally, the magnitude of the difference in whether the employer sent an email and requested the resume to be resent is 68 percent (19 percent for minority and 31 percent for majority applicants).

Next, we explore the effect of signaling recent unemployment on attention, i.e. the effect of an unambiguously negative signal about the applicant’s quality. We find that employers consistently adjust their attention based on such information. Compared to the baseline condition with no information about unemployment history, indicating an 18-months-unemployment significantly reduces all three measures of information-acquisition effort (Table 3.6 and Panel A of Table 3.7).

Observation 7: A signal of an applicant’s lower quality, observed by an employer prior to reading a resume reduces an employers’ attention to the resume.

Similarly as in the case of name effects, we find the magnitude of the gap in attention increases at higher levels of effort to acquire a resume. The difference between the baseline condition and the 18-month unemployment condition is 13 percent for the likelihood of opening a resume, 23 percent for the number of attempts to open a resume, and 33

percent for the likelihood of requesting the resume to be re-sent. The effect of the 2-month unemployment condition is generally smaller in size and less significant statistically (Panel A of Table 3.7). Specifically, the difference between the 2-month and 18-month unemployment conditions is statistically significant for the number of attempts to open a resume and for the likelihood of requesting the resume to be re-sent (Panel B).

Last, we do not find evidence of a systematic interaction effect of minority names and unemployment conditions on attention (Columns 2, 4 and 6, Panel A of Table 3.7). In other words, the 18-month unemployment condition lowers an employer's effort to read a resume for both minority as well as majority applicants. Also, in the baseline condition, which is most comparable to the labor market experiment in the Czech Republic, the minority name significantly lowers the number of clicks on the resume as well as the likelihood of requesting the resume to be re-sent. The negative effect is small and statistically insignificant for the likelihood of opening a resume.

3.6 Links to Theories

We now consider which models can explain the set of findings from the three correspondence tests. Although it is likely that the observed discrimination in terms of invitation rates arises, at least in part, due to reasons highlighted in standard economic models of discrimination—preference-based and statistical discrimination models—, these models cannot explain the complete set of findings, in particular the observed discrimination in attention, an important input for selection decision.

In purely preference-based models of discrimination, individuals do not discriminate due to lack of information and thus imperfect information and attention do not enter the model. There are several models that generate discrimination via imperfect information. Their common feature is that all observable actions prior to selection decisions are the same and discrimination arises at the moment of selection decision when the imperfect information is used. In the first class of statistical discrimination models decision makers take into account observable individual characteristics, while using an observable group attribute, such as ethnicity, to proxy unobservable individual quality (Phelps 1972; Arrow 1973).

The second class of statistical discrimination models emphasizes a lower precision of observable signals as a source of discrimination (Aigner and Cain 1977). Specifically, signals about individuals that economic agents receive are more precise for majority ap-

plicants compared to minority applicants, perhaps due to cultural dissimilarity (Cornell and Welch 1996). Thus, the difference in the precision of information about individuals across ethnic groups is assumed, i.e. is exogenous, and is not due to differences in efforts to acquire information. In contrast to these models, the experiments reveal that discrimination begins earlier, already during the information-acquisition stage, creating differences in information imperfection across groups at the moment when agents finally make decisions.

Do decision-makers allocate costly attention endogenously? To assess that, we consider whether it can explain the main findings: (i) In the labor market employers pay more attention to majority compared to minority candidates, while in the rental housing market landlords pay more attention to minority compared to majority candidates. (ii) The gap in resume acquisition is greater when acquiring a resume requires writing an email requesting re-sending a resume compared to simple clicking on a hyperlink. (iii) Signaling recent unemployment—another type of negative signal besides a minority name—lowers attention to an applicant on the labor market, similarly as minority name does.

First, the model predicts a switch in relative attention if markets differ in selectivity. In the labor market, where selectivity is high—since firms select only a few top applicants for an interview—the expected benefits from reading a resume are smallest for the a priori least attractive group, while the benefits of inspection are greater for this group in the housing market, where the overall invitation rate is high. Second, it predicts that at low costs of information acquisition the increased cost increases the motivation of the decision-maker to optimize attention and thus the gap in resume acquisition is predicted to increase when an employer needs to write an email compared to clicking on a hyperlink in order to get a resume. Third, it implies that any signal of an applicant's quality should affect attention to subsequent information, independently of whether the signal concerns ethnic status or some other characteristic relevant for beliefs about quality, such as signaling recent unemployment. Thus, we conclude the predictions of the model are consistent with all three empirical findings.

Note that the switch in relative attention across markets is predicted to arise if decision makers have racist preferences, believe that minority candidates are of lower quality on average, or expect members of a minority group to be more alike. Therefore, based on documenting the switch across markets, we still cannot distinguish between these sources. Indicative evidence on this question is provided by two supplementary online surveys, in

which we find that landlords as well as employers expect to be less satisfied with minority candidates, indicating either dislike or belief about a lower objective quality. At the same time, we find virtually no differences in the variance of expected satisfaction across ethnic groups.

It should be noted that the selection of tenants by landlords differs from the hiring decisions of employers in many ways, and thus attributing the switching results to differences in selectivity needs to be taken cautiously. For example, desired applicant's qualities may differ across the markets—landlords may be concerned about a tenant's ability to reliably pay rent and not cause property damage, while employers may focus on the type of education and qualifications relevant for a given job. If minority applicants were considered a priori better tenants and worse employees, then such a combination of beliefs could, in principle, explain the observed switch. However, this interpretation is inconsistent with the observed discrimination in selection decisions as well as results of online surveys, which suggest that minorities are negatively stereotyped in both markets.

Similarly, the expectation of the relative precision of available signals (Aigner and Cain 1977; Cornell and Welch 1996) may also differ across markets. Our model predicts more attention to groups with more precise signals since the benefit of paying a unit of attention to such groups is higher. The switch in relative attention could then be explained if the expected precision of the available signals also switched, i.e. if employers expected signals about minorities to be less informative, while landlords expected the same about majority applicants. We asked employers and landlords about these types of expectations in online surveys. Their responses support the former but not the latter.

Next, it is also possible that knowing more about minority applicants may be more important for a landlord than for an HR manager, perhaps because landlords may be more likely to interact intensively with tenants than HR managers with employees. However, the importance of the decision is predicted to affect overall levels of attention and thus influence the magnitude of attention discrimination, but it is not predicted to lead to a switch in relative attention across groups. Potentially, there might be other differences across markets, which could explain the switch, although the explanation based on differences in selectivity seems to be the most parsimonious.

While we propose a model in which decision makers consciously allocate attention based on expected benefits in each instance, the observed attention choices of employers and landlords in the experiments may be conscious as well as automatic based on simplifying screening heuristics. Our model can help to explain why discrimination heuristics may

arise, for example by trial and error or by an initial conscious setting of screening rules that prove to work reasonably well and are later used automatically. Nevertheless, attention allocation can also be affected by unconscious mental associations against negatively-stereotyped groups, termed implicit discrimination (Bertrand, Chugh, and Mullainathan 2005) and supported by intriguing evidence from Implicit Association Tests (Greenwald, McGhee, and Schwartz 1998; Stanley, Phelps, and Banaji 2008). Such unconscious biases, which may operate in parallel with—and sometimes in contradiction of—one’s conscious intentions, could also explain the lower observed attention to minorities on labor markets.

3.7 Concluding Remarks

One of the main insights from information economics is that even very small frictions in information acquisition can have large effects on economic outcomes (Diamond 1971; Sims 2003). At the same time, imperfect information is central to explaining discrimination in markets since the seminal work of Phelps (1972) and Arrow (1973). Yet, there is no theory or direct evidence studying how the small costs of information acquisition may create differences in the form of imperfect information about individuals based on their observable group attributes. This is what we provide.

We first describe how choices of attention affect discrimination in theory. We show that if attention is costly, prior beliefs about ethnic groups enter the final decision not only through Bayesian updating, as in the standard model of statistical discrimination, but also earlier through the choice of attention to available information ("attention discrimination"). As a result, prior beliefs have the potential for a larger impact on discrimination (in most types of markets) and discrimination in the selection of applicants can arise even when the decision makers have the same preferences across different groups, when all the relevant information is available, and when obtaining information about different groups is equally difficult. Costly attention can similarly magnify the role of animus.

In the empirical part, we identify attention discrimination in practice. We develop new tools for field experiments using the Internet, and monitor information acquisition by employers and landlords about applicants prior to a selection decision for a job interview and an apartment viewing. A set of three experiments in two countries reveals that signals of an applicant’s minority status systematically affect attention to easily available information about the applicant (e.g., resume). In line with the model, the observed patterns of attention allocation are consistent with economic agents considering reading

applications for a job or apartment as a costly activity and choosing the level of attention with an eye for expected benefits of reading, taking into account an applicant's observable group attribute and desired level of quality in a given market. We also discuss alternative interpretations, in particular the potential role of an unconscious bias in attention.

The key insight that willingness to process information at hand represents an additional barrier for applicants with unfavorable group attributes points towards several promising directions for future research as well as thoughts about policy, and we mention a few. The model implies the important role of the timing of when a group attribute is revealed—the later a decision maker learns a group attribute, such as name, the smaller the asymmetry in attention to subsequent information such as education or qualification. It is intriguing that employers in the public as well as private sector have recently started to introduce name-blind resumes,⁴⁰ in part because researchers produced evidence indicating that blind auditioning (Goldin and Rouse 2000) and name-blind resumes (Skans and Åslund 2012) can reduce discrimination. Understanding practical implications of attention discrimination and which policies may be the most appropriate to attenuate it, without imposing too many restrictions on a firm's choices, is an important area to explore. The idea that early signals have disproportionately large effects on outcomes also has implications for members of negatively stereotyped groups who cannot take for granted that employers will learn their qualities when reading a resume. It might help to provide positive signals early on, for example, by mentioning previous relevant job experience already in the introductory email. Although such information does not reveal anything new since it is fully contained in the resume, it may prevent the employers from putting the resume aside.

Next, the lower predicted and observed attention to negatively stereotyped groups in selective markets can help explain why African-Americans and minorities were found to face lower returns to higher quality resumes in the labor markets in the US and Sweden, respectively (Bertrand and Mullainathan 2004; Bursell 2007). Based on this, we speculate that in the long-term endogenous attention lowers incentives of negatively stereotyped groups to acquire human capital in the first place, and could make beliefs about differences in quality potentially self-fulfilling. Last, if the effect of recent unemployment on the attention of employers is similar when unemployment is signaled in the introductory

⁴⁰Name-blind resumes have recently been implemented for hiring workers in the public sector in Belgium, the Netherlands, and Sweden. The policy is being piloted in Germany among several major companies, including Deutsche Post, Deutsche Telekom, L'Oréal, and Procter & Gamble.

email, as in our experiment, just as when it is reported in the resume, then endogenous attention may also contribute to greater long-term unemployment.

In addition to presenting novel empirical findings, the experimental design distinguishes itself by offering a methodological contribution. Our analysis joins efforts in laboratory settings to test decision-making processes with enhanced measurement tools, in particular by monitoring information acquisition (Camerer et al. 1993; Costa-Gomes, Crawford, and Broseta 2001; Gabaix, Laibson, and Moloche 2006). We show that the widespread use of the Internet by economic decision makers opens the possibility of collecting "process data" as a part of a natural field experiment as well. By this, researchers can study in greater detail the processes taking place inside the "black box" and can better inform theories and policy-makers on issues, including those that are sensitive and hard to study in the laboratory (Levitt and List 2007), of which discrimination is one important example.

Table 3.1: Czech Rental Housing Market — Invitation Rates and Information Acquisition by Ethnicity, Comparison of Means

	White majority name (W) (1)	Pooled Asian and Roma minority name (E) (2)	p.p. difference: W-E, (p-value) (3)	Asian minority name (A) (4)	p.p. difference: W-A, (p-value) (5)	Roma minority name (R) (6)	p.p. difference: W-R, (p-value) (7)	p.p. difference: R-A, (p-value) (8)
Panel A: Invitation for a flat visit								
No Information Treatment (n=451)	0.78	0.41	37 (0.00)	0.39	39 (0.00)	0.43	36 (0.00)	3 (0.57)
Monitored Information Treatment (n=762)	0.72	0.49	23 (0.00)	0.49	23 (0.00)	0.49	23 (0.00)	0 (0.92)
Monitored Information Treatment ^a (n=293)	0.84	0.66	18 (0.00)	0.71	13 (0.00)	0.62	21 (0.00)	-9 (0.20)
Monitored Information Treatment ^b (n=469)	0.66	0.37	29 (0.00)	0.35	31 (0.00)	0.39	27 (0.00)	4 (0.51)
Treatment with additional text in the email (n=587)	0.78	0.52	26 (0.00)	0.49	29 (0.00)	0.55	23 (0.00)	5 (0.29)
Panel B: Information acquisition in the Monitored Information Treatment								
Opening applicant's personal website	0.33	0.41	-8 (0.03)	0.38	-5 (0.24)	0.44	-11 (0.01)	6 (0.15)
Number of pieces of information acquired	1.29	1.75	-0.46 (0.01)	1.61	-0.32 (0.09)	1.88	-0.59 (0.00)	0.27 (0.17)
At least one piece of information acquired	0.30	0.40	-10 (0.01)	0.37	-7 (0.12)	0.44	-13 (0.00)	7 (0.12)
All pieces of information acquired	0.19	0.26	-8 (0.02)	0.24	-6 (0.12)	0.28	-10 (0.01)	4 (0.33)
Number of pieces of information acquired ^a	3.91	4.24	-0.33 (0.06)	4.23	-0.32 (0.15)	4.25	-0.34 (0.09)	0.02 (0.90)
At least one piece of information acquired ^a	0.92	0.98	-6 (0.02)	0.97	-5 (0.15)	0.98	-7 (0.03)	2 (0.47)
All pieces of information acquired ^a	0.56	0.64	-7 (0.23)	0.64	-8 (0.30)	0.64	-7 (0.30)	-0 (0.96)

Notes: Means. Panel A reports how name affects invitation for a flat visit and Panel B how it affects information acquisition in the Monitored Information Treatment. Columns 3, 5, 7 and 8 report differences in percentage points, in the parentheses we report p-value for a t-test testing the null hypothesis that the difference is zero. The differences in the number of pieces of information acquired on the website are reported in absolute terms, not in percentage points.

^a The numbers are reported for the sub-sample of landlords who opened an applicant's website. ^b The numbers are reported for the sub-sample of landlords who did not open an applicant's website.

Table 3.2: Czech Rental Housing Market — Invitation Rates by Ethnicity, Regression Analysis

Dependent variable	Invitation for an Apartment Viewing				
	(1)	(2)	(3)	(4)	(5)
Panel A					
Sample:	No Information Treatment		Whole sample		
			All	W majority name	E minority name
Ethnic minority name	-0.39*** (0.04)		-0.37*** (0.04)		
Asian minority name		-0.41*** (0.05)			
Roma minority name		-0.39*** (0.05)			
Monitored Information Treatment			-0.08 (0.06)	-0.06 (0.05)	0.08** (0.04)
Ethnic minority name*Monitored Information Treatment			0.16** (0.06)		
Additional text in the email - with high school			-0.00 (0.07)	-0.00 (0.06)	0.08* (0.05)
Ethnic minority name*Additional text in the email - with high school			0.08 (0.08)		
Additional text in the email - with college			0.00 (0.07)	0.01 (0.06)	0.15*** (0.05)
Ethnic minority name*Additional text in the email - with college			0.14* (0.08)		
Observations	451	451	1,800	599	1,194
Panel B					
Sample:	Monitored Information Treatment				
				Landlords who opened applicant's website	Landlords who did not open applicant's website
	All	All	All		
Ethnic minority name	-0.23*** (0.04)	-0.26*** (0.04)	-0.28*** (0.04)	-0.18*** (0.05)	-0.29*** (0.05)
Opening applicant's website		0.26*** (0.04)	0.21*** (0.07)		
Ethnic minority name*Opening applicant's website			0.07 (0.09)		
Observations	762	762	762	293	469

Notes: Probit, marginal effects (dF/dx) in all Columns of both Panels. Robust standard errors in parentheses. In Columns 1-3 of Panel A and all Columns of Panel B the omitted variable is a White majority name. In Columns 4-5 of Panel A the omitted variable is a dummy for No Information Treatment. In all Columns of both Panels, we control for a dummy variable indicating a landlord being a female, a dummy variable indicating an unknown gender of a landlord (the mean of this variable is 0.02), size of an apartment, price of an apartment rental, and a dummy variable indicating an equipped apartment.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 3.3: Czech Rental Housing Market — Information Acquisition by Ethnicity, Regression Analysis

Dependent variable:	Opening applicant's personal website		Number of pieces of information acquired		Number of pieces of information acquired about education and occupation		Number of pieces of information acquired about personal characteristics	
Sample:	Monitored Information Treatment - all observations		Monitored Information Treatment - sub-sample of landlords who opened applicant's website					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ethnic minority name	0.08** (0.04)		0.46*** (0.16)		0.18** (0.09)		0.12 (0.13)	
Asian minority name		0.05 (0.04)		0.31 (0.19)		0.17* (0.10)		0.07 (0.14)
Roma minority name		0.11*** (0.04)		0.60*** (0.19)		0.18* (0.09)		0.15 (0.14)
Constant			1.42*** (0.34)	1.42*** (0.34)	1.70*** (0.16)	1.71*** (0.16)	2.46*** (0.21)	2.47*** (0.21)
Observations	762	762	762	762	293	293	293	293

Notes: Probit, marginal effects (dF/dx) in Columns 1 and 2. OLS in Columns 3-8. Robust standard errors in parentheses. In Columns 3-4, the dependent variable is number of pieces of information a landlord viewed on applicant's personal website - minimum is 0 and maximum is 5. In Columns 5-6 it is a number of pieces of information about education and occupation he/she uncovered - minimum is 0 and maximum is 2, and in Columns 7-8 it is a number of pieces of information about personal characteristics (age, smoking habits, marital status) he/she uncovered - minimum is 0 and maximum is 3. In all Columns the omitted variable is a White majority name. In all Columns, we control for a dummy variable indicating a landlord being a female, a dummy variable indicating an unknown gender of a landlord, size of an apartment, price of an apartment rental, and a dummy variable indicating an equipped apartment.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level

Table 3.4: Czech Labor Market — Invitation Rates and Information Acquisition by Ethnicity, Comparison of Means

	White majority name (W) (1)	Pooled Asian and Roma minority name (E) (2)	p.p. difference: W-E, (p-value) (3)	Asian minority name (A) (4)	p.p. difference: W-A, (p-value) (5)	Roma minority name (R) (6)	p.p. difference: W-R, (p-value) (7)	p.p. difference: R-A, (p-value) (8)
Panel A: Employer's response								
Callback	0.43	0.20	23 (0.00)	0.17	26 (0.00)	0.25	18 (0.01)	8 (0.22)
Invitation for a job interview	0.14	0.06	8 (0.03)	0.05	9 (0.03)	0.08	6 (0.18)	3 (0.46)
Invitation for a job interview ^a	0.19	0.09	10 (0.06)	0.09	10 (0.12)	0.10	9 (0.16)	1 (0.83)
Panel B: Information acquisition								
Opening applicant's resume	0.63	0.56	7 (0.22)	0.47	16 (0.03)	0.66	-3 (0.69)	19 (0.01)
Acquiring more information about qualification ^a	0.16	0.10	6 (0.27)	0.06	10 (0.12)	0.14	2 (0.73)	8 (0.24)
Acquiring more information about other characteristics ^a	0.18	0.18	0 (0.92)	0.19	-1 (0.85)	0.18	0 (0.99)	1 (0.85)

Notes: Means. Panel A reports how name affects callback and invitation for a job interview and Panel B how it affects information acquisition. Columns 3, 5, 7 and 8 report differences in percentage points, in the parentheses we report p-value for a t-test testing the null hypothesis that the difference is zero. Acquiring more information about qualification is a dummy variable indicating whether an employer clicked on "learn more" buttons on a resume to acquire more information about education, experience, and skills. Acquiring more information about other characteristics is a dummy variable indicating whether she/he acquired more information about hobbies and contact information.

^aThe numbers are reported for the sub-sample of employers who opened applicant's resume.

Table 3.5: Czech Labor Market — Invitation Rates and Information Acquisition by Ethnicity, Regression Analysis

Dependent variable	Invitation for a job interview		Opening applicant's resume		Acquiring more information about qualification		Acquiring more information about other characteristics	
Sample:	All		All		Employers who open applicant's resume			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ethnic minority name	-0.09*** (0.04)		-0.08 (0.06)		-0.07 (0.06)		-0.01 (0.064)	
Asian minority name		-0.08** (0.03)		-0.16** (0.07)		-0.10* (0.05)		-0.00 (0.08)
Roma minority name		-0.06* (0.03)		0.03 (0.08)		-0.03 (0.06)		-0.02 (0.07)
Required high school	-0.03 (0.051)	-0.02 (0.049)	0.10 (0.095)	0.11 (0.097)	0.08 (0.056)	0.08 (0.054)	0.01 (0.097)	0.01 (0.097)
Required experience	-0.09*** (0.026)	-0.09*** (0.026)	-0.00 (0.069)	0.01 (0.069)	-0.03 (0.057)	-0.03 (0.055)	-0.09 (0.064)	-0.09 (0.064)
Sector of sales and services	-0.02 (0.037)	-0.02 (0.037)	-0.05 (0.067)	-0.04 (0.068)	-0.00 (0.057)	-0.00 (0.057)	0.01 (0.068)	0.01 (0.067)
Application in holiday period	0.03 (0.037)	0.02 (0.037)	0.04 (0.067)	0.04 (0.068)	0.01 (0.058)	0.02 (0.058)	0.02 (0.069)	0.02 (0.069)
Observations	274	274	274	274	160	160	160	160

Notes: Probit, marginal effects (dF/dx), robust standard errors in parentheses. In Columns 5-6, the dependent variable is a dummy variable indicating whether an employer clicked on "learn more" buttons on a resume to acquire more information about education, experience, and skills; in Columns 7-8 it indicates whether she/he acquired more information about hobbies and contact information. In all Columns the omitted variable is a White majority name and we control for dummy variables indicating required high school education, required previous experience, application being sent during a holiday period (August), and application in the sector of sales and services.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 3.6: German Labor Market — Information Acquisition by Ethnicity, Comparison of Means

	(1)	(2)	(3)	(4)	(5)
Panel A: Effect of name on information acquisition					
	White majority name (W) (n=366)	Ethnic minority name (E) (n=379)	difference: W-E (p-value)		
Opening applicant's resume	0.81	0.75	0.06 (0.06)		
Number of attempts to open applicant's resume	2.81	2.14	0.67 (0.00)		
Email request to re-send resume	0.31	0.19	0.13 (0.00)		
Panel B: Effect of signal about unemployment on information acquisition					
	No information (N) (n=372)	2 months unemployed (2M) (n=187)	18 months unemployed (18M) (n=186)	p.p. difference N-2M (p-value)	p.p. difference N-18M (p-value)
Opening applicant's resume	0.83	0.73	0.73	10 (0.01)	10 (0.00)
Number of attempts to open applicant's resume	2.66	2.51	2.04	0.15 (0.56)	0.63 (0.01)
Email request to re-send resume	0.27	0.26	0.18	2 (0.66)	9 (0.02)

Notes: Means. Panel A reports how information acquisition is affected by name and Panel B how it is affected by the signal about recent unemployment. In Column 3 of Panel A and Columns 4-5 of Panel B we report differences in means between White majority and ethnic minority group, in the parentheses we report p-value for a t-test testing the null hypothesis that the difference is zero. The differences in the number of attempts to open applicant's resume are reported in absolute terms, not in percentage points.

Table 3.7: German Labor Market — Information Acquisition by Ethnicity, Regression Analysis

Dependent variable	Opening applicant's resume		Number of attempts to open applicant's resume		Email request to re-send resume	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A			Sample: All			
Ethnic minority name	-0.06* (0.03)	-0.02 (0.05)	-0.68*** (0.20)	-0.82*** (0.28)	-0.15*** (0.03)	-0.14*** (0.04)
2 months unemployed	-0.10** (0.04)	-0.04 (0.06)	-0.10 (0.24)	-0.09 (0.35)	-0.01 (0.04)	0.02 (0.05)
18 months unemployed	-0.12*** (0.04)	-0.10* (0.06)	-0.64*** (0.24)	-0.92*** (0.34)	-0.07** (0.03)	-0.08* (0.04)
Ethnic minority name*2 months unemployed		-0.11 (0.09)		-0.01 (0.48)		-0.06 (0.06)
Ethnic minority name*18 months unemployed		-0.03 (0.08)		0.57 (0.49)		0.01 (0.08)
Constant			3.17*** (0.27)	3.24*** (0.29)		
Observations	745	745	745	745	745	745
Panel B			Sample: 2 months unemployed and 18 months unemployed			
Ethnic minority name	-0.10** (0.05)	-0.15** (0.07)	-0.51* (0.29)	-0.85** (0.40)	-0.16*** (0.04)	-0.19*** (0.05)
18 months unemployed	-0.01 (0.05)	-0.06 (0.07)	-0.55* (0.28)	-0.91** (0.41)	-0.07* (0.04)	-0.10** (0.05)
Ethnic minority name*18 months unemployed		0.09 (0.08)		0.70 (0.57)		0.09 (0.09)
Constant			3.14*** (0.45)	3.33*** (0.48)		
Observations	373	373	373	373	373	373

Notes: Probit, marginal effects (dF/dx) in Columns 1-2 and 5-6, OLS in Columns 3-4. Robust standard errors in parentheses. In all Columns of both Panels, the omitted variable is a White majority name and we control for dummy variables indicating required high school education, required previous experience, position in a city with more than million inhabitants, application being sent in holiday period (August), and a set of four dummy variables indicating the sector (manufacturing and construction, information and communication, administration, and professional, scientific and technical activities). In all Columns of Panel, the omitted variable is No information about unemployment, while in all Columns of Panel B it is 2 months unemployed.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

3.A Appendix 3

3.A.1 Supplementary Material to Section 3.2

Proof of Proposition 1:

(A) First, let us consider the effect of a decrease in q_G in a cherry-picking market. In the first stage, the DM chooses between rejecting the applicant, inviting him or her, and acquiring information about q_1 . By definition, in a cherry picking market $\text{payoff}(\text{reject}) > \text{payoff}(\text{invite})$, and thus the choice is between rejecting the applicant right away and acquiring more information first. A decrease in q_G does not affect $\text{payoff}(\text{reject})$, but it decreases $\text{payoff}(\text{info})$. Therefore, such a change weakly decreases attention. Obviously, an increase in d_G or C_2 has the same effect. Similarly, a decrease in σ_G^2 does not affect $\text{payoff}(\text{reject})$, while it decreases $\text{payoff}(\text{info})$, too. This is because the payoff from invitation to the second stage, see Definition on page 130 in the main text,

$$E[\max(R, E[\max(R, q - d_G)|q_1] - C_2)]$$

is due to the max-operators increasing in a mean-preserving spread; positive changes count toward a higher payoff, while some of the negative changes are filtered out by the reservation payoff R . If σ_G^2 increases, then $E[\max(R, q - d_G)|q_1] - C_2$ increases for all q_1 , and thus the distribution in Figure 3.1 shifts to the right in the sense of first-order stochastic dominance.

In the lemon-dropping market, the situation is only slightly more complicated. Now, $\text{payoff}(\text{invite}) > \text{payoff}(\text{reject})$. For each realization of q_1 the corresponding positive impact of an increase in q_G is at least as high for $\text{payoff}(\text{invite})$ as for $\text{payoff}(\text{info})$. See the payoffs on page 130 in the main text - any increase in $E[\max(R, q - d_G)|q_1]$ contributes directly to an increase in $\text{payoff}(\text{invite})$, while it contributes to an increase in $\text{payoff}(\text{info})$ only when $E[\max(R, q - d_G)|q_1]$ is higher than $(R + C_2)$. Therefore, as a result of an increased q_G , $\text{payoff}(\text{invite})$ increases more than $\text{payoff}(\text{info})$, which decreases attention. Arguments showing the stated effects of changes in d_G , C_2 , and σ_G^2 are completely analogous.

(B) A higher C_1 decreases $\text{payoff}(\text{info})$, while leaving $\text{payoff}(\text{invite})$ and $\text{payoff}(\text{reject})$ unchanged, and thus it weakly decreases attention in either market. Similarly, a lower $\sigma_{G,1}^2$ does not affect $\text{payoff}(\text{reject})$ or $\text{payoff}(\text{invite})$, but it decreases $\text{payoff}(\text{info})$,

and thus it leads to a weakly lower attention.

QED.

Proof of Corollary 1:

(A) First, in a cherry-picking market, attention is a necessary prerequisite for being accepted in the selection decision, since the applicant is rejected if no additional information is acquired. Higher information acquisition in the first stage thus weakly increases the probability that the applicant is invited to the second stage. Next, the applicant's quality is observed upon invitation to the second stage, and thus acceptance in the second stage, conditional on being invited to the second stage, is for a given applicant independent of information acquisition in the first stage. Therefore, higher information acquisition in the first stage weakly increases the probability of acceptance in the second stage. Implications of information in the lemon-dropping market are analogous, and the remaining step connecting group characteristics and implications of endogenous attention is an immediate implication of Proposition 1.

(B) The steps are completely the same as in (A) except for the fact that Proposition 1 states that a group dissimilarity decreases attention in either market.

QED.

3.A.2 Supplementary Material to Sections 3.3-3.5

Wording of Application Email—Czech Rental Housing Market

1. "Dear Sir/Madam, I am writing because I am very interested in renting the apartment that you have advertised. When would be a good time to come see the apartment? Best regards, Phan Quyet Nguyen"
2. Adding a link to personal website: "Dear Sir/Madam, I am writing because I am very interested in renting the apartment that you have advertised. When would be a good time to come see the apartment? Best regards, Phan Quyet Nguyen, phan.quiet.nguyen.sweb.cz"
3. Adding a sentence with applicant's characteristics: "Dear Sir/Madam, I am writing because I am very interested in renting the apartment that you have advertised. I am a thirty-year-old man, I am single, I have a college [a high-school] degree,

and I do not smoke. I have a steady job (with a regular paycheck) at a company. When would be a good time to come see the apartment? Best regards, Phan Quyet Nguyen"

Wording of Application Email—Czech Labor Market

1. "Dear Sir/Madam, I am writing because I am very interested in the Real Estate Agent job position advertised by your company. You can find my resume in this hyperlink: phanquyetnguyen1982.sweb.cz. Best regards, Phan Quyet Nguyen"

Wording of Application Email—German Labor Market

1. "Dear Sir/Madam, I am writing because I am very interested in the Real Estate Agent job position advertised by your company. You can find my resume in this hyperlink: fatihyildiz1982.webege.com. Best regards, Fatih Yildiz"
2. Adding information about 2 months unemployment: "Dear Sir/Madam, I have been searching for a job for two months and I am writing because I am very interested in the Real Estate Agent job position advertised by your company. You can find my resume in this hyperlink: fatihyildiz1982.webege.com. Best regards, Fatih Yildiz"
3. Adding information about 18 months unemployment: "Dear Sir/Madam, I have been searching for a job for a year and half and I am writing because I am very interested in the Real Estate Agent job position advertised by your company. You can find my resume in this hyperlink: fatihyildiz1982.webege.com. Best regards, Fatih Yildiz"

Figure 3.A.1: Applicant's Personal Website Snapshot, Czech Rental Housing Market

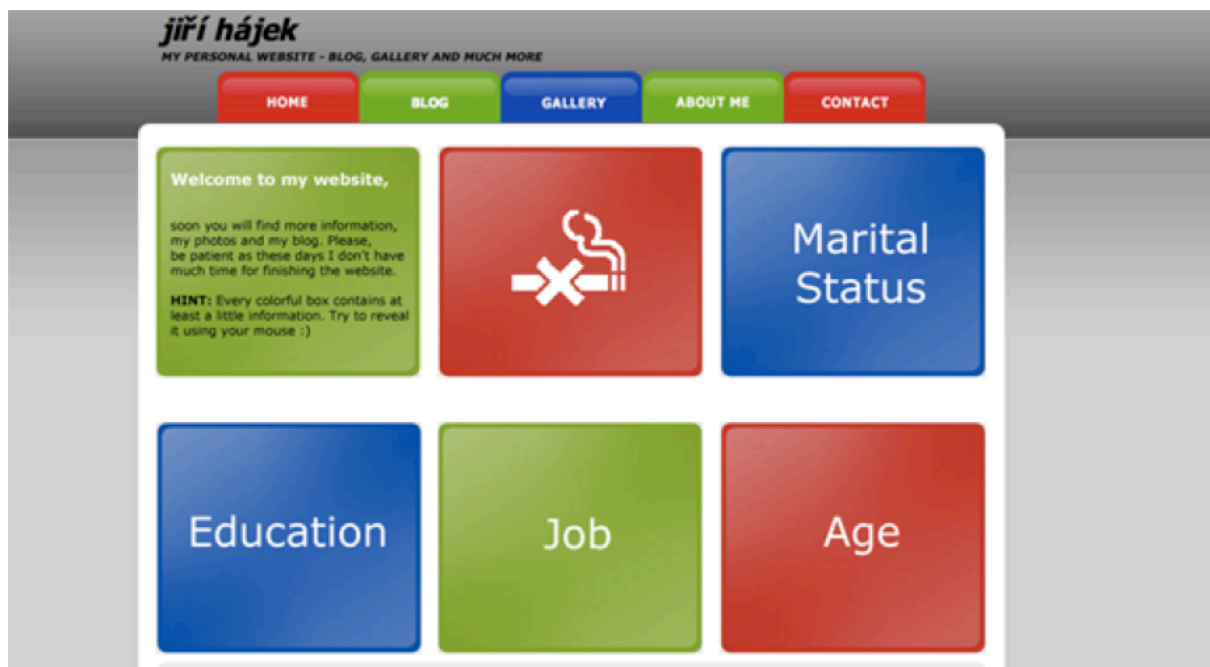


Figure 3.A.2: Applicant's Online Resume, Czech Labor Market

Left Part: A Snapshot After Opening the Website (a Shorter Form)

Right Part: A Snapshot After Expanding Education and Experience Categories

PHAN QUYET NGUYEN CURRICULUM VITAE		phanquyetnguyen1982@seznam.cz (+420) 605 174 397 [more]
Education [more]	BUSINESS ACADEMY PRAGUE 6, KRUPKOVO NÁMĚSTÍ	1997-2001
Experience [more]	AZPIRO, LTD. Administrative support of consultants, PC work	2006-2010
	AUTO NELLY LTD. International purchasing assistant	2001-2005
	MULTIMEDIA MED, LTD. Market research; customer surveys	1999-2000
Skills [more]	Language skills English language: Fluent, passed final exam from German language: Intermediate.	
	Driving licence Type B	

PHAN QUYET NGUYEN CURRICULUM VITAE		phanquyetnguyen1982@seznam.cz (+420) 605 174 397 [less] Marital status: Single Date of birth: July 13th, 1982
Education [less]	BUSINESS ACADEMY PRAGUE 6, KRUPKOVO NÁMĚSTÍ	1997-2001
	Final exam grades: Accounting - B Economics - A Set of vocational courses - A English language - B	
	Subjects studied: Written and electronic communication, accounting, economics, statistics, tourism management, English and German	
Experience [less]	AZPIRO, LTD. Administrative support of consultants, PC work	2006-2010
	Document management; administrative support of consultants; PC work mainly with Microsoft Excel and Access; creating client databases with information about projects, project content, costs and price lists. For references see References section .	
	AUTO NELLY LTD. International purchasing assistant	2001-2005
	Assistance with purchases; communication with international customers; PC work, mainly with Microsoft Word and Excel on client management, purchases and price databases.	

Table 3.A.1: Czech Rental Housing Market - Design of the Experiment

	No Information Treatment	Monitored Information Treatment		Treatment with additional text in the email Email: name, info about education, occupation, age, marital status, smoking	
	Email: name	Email: name and hyperlink to website Website: info about education, occupation, age, marital status, smoking			
		High school degree	College degree	High school degree	College degree
White majority name	X	X	X	X	X
Asian minority name	X	X	X	X	X
Roma minority name	X	X	X	X	X

Table 3.A.2: Survey III - Comparison of the Names Used in the Czech Experiments

Dependent variable	Education level		Quality of housing			
	High school (1)	University (2)	Lodging (3)	Rented flat (4)	Own flat (5)	Own house (6)
Panel A: Comparison of all three names (omitted majority-sounding name)						
Roma-sounding name	-1.82*** (0.24)	-2.08*** (0.26)	2.45*** (0.25)	0.19 (0.22)	-1.37*** (0.23)	-1.31*** (0.25)
Asian-sounding name	-0.61** (0.24)	-0.39 (0.25)	0.70*** (0.24)	-0.24 (0.21)	-0.53** (0.23)	-0.16 (0.24)
Constant	5.06*** (0.17)	3.71*** (0.18)	1.63*** (0.18)	4.42*** (0.15)	4.00*** (0.17)	3.14*** (0.17)
Observations	246	246	246	245	246	246
Panel B: Comparison of minority-sounding names (omitted Asian-sounding name)						
Roma-sounding name	-1.21*** (0.25)	-1.68*** (0.26)	1.75*** (0.26)	0.44** (0.22)	-0.84*** (0.24)	-1.15*** (0.24)
Constant	4.45*** (0.17)	3.31*** (0.18)	2.34*** (0.18)	4.17*** (0.15)	3.47*** (0.16)	2.98*** (0.17)
Observations	167	167	167	166	167	167

OLS in all Columns of all Panels. Standard errors in parentheses. Majority-sounding name is Jiri Hajek, Roma-sounding name is Gejza Horvath and Asian-sounding name is Phan Quyet Nguyen. The dependent variables are measured on a scale 0-7. 0 means that a respondent considered it impossible for a person with the given name to have high school (university) education and to live in lodging (in a rented flat, in an own flat, in an own house). 7 means that a respondent considered it certain.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 3.A.3: Czech Rental Housing Market - Randomization Check

Experimental manipulation:	Name of applicant						Access to information		
	White majority name (1)	Ethnic minority name (2)	t-test p-value (3)	Asian minority name (4)	Roma minority name (5)	F-stat p-value (6)	No Information (7)	Monitored Information (8)	t-test p-value (9)
Female landlord	0.46 (0.50)	0.49 (0.50)	0.26	0.46 (0.50)	0.51 (0.50)	0.15	0.50 (0.50)	0.48 (0.50)	0.49
Size of the apartment (hundreds of m ²)	0.47 (0.15)	0.47 (0.14)	0.83	0.47 (0.15)	0.47 (0.14)	0.96	0.47 (0.15)	0.47 (0.14)	0.79
Price of the apartment (ths. CZK)	9.03 (2.94)	8.89 (2.86)	0.33	8.79 (2.82)	8.98 (2.89)	0.32	8.87 (2.95)	8.96 (2.98)	0.60
Apartment equipped	0.15 (0.36)	0.16 (0.37)	0.53	0.16 (0.37)	0.17 (0.38)	0.74	0.13 (0.34)	0.17 (0.37)	0.12
Observations	606	1194		569	625		451	762	

Notes: Means. Standard deviations in parentheses. Column 3 reports p-value for a t-test testing the null hypothesis that the means are equal for applicants with a majority-sounding name and a minority-sounding name (Asian and Roma minority pooled together). Column 6 reports p-value for an F-test testing the null hypothesis that the means are equal across all three groups of applicants. Column 9 reports p-value for an F-test testing the null hypothesis that the means are equal in the No Information Treatment and in the Monitored Information Treatment.

Table 3.A.4: Czech Rental Housing Market - Callback by Ethnicity

	White majority name (W) (1)	Pooled Asian and Roma minority name (E) (2)	p.p. difference: W-E, (p-value) (3)	Asian minority name (A) (4)	p.p. difference: W-A, (p-value) (5)	Roma minority name (R) (6)	p.p. difference: W-R, (p-value) (7)	p.p. difference: R-A, (p-value) (8)
No Information Treatment	0.89	0.58	32 (0.00)	0.54	35 (0.00)	0.61	28 (0.00)	7 (0.19)
Monitored Information Treatment	0.81	0.62	19 (0.00)	0.60	21 (0.00)	0.63	18 (0.00)	3 (0.49)
Monitored Information Treatment ^a	0.89	0.75	15 (0.00)	0.75	14 (0.01)	0.74	15 (0.01)	0 (0.89)

Notes: Means. The table reports the likelihood of callback across names and treatments. Columns 3, 5, 7 and 8 report differences in percentage points, in the parentheses we report p-value for a t-test testing the null hypothesis that the difference is zero. ^a The numbers are reported for the sub-sample of landlords who opened applicant's website.

Table 3.A.5: Czech Rental Housing Market - Information Search

	All (1)	White majority name (W) (2)	Ethnic minority name (E) (3)	p.p. difference: W-E, (p-value) (4)
Panel A				
Sample:		Monitored information treatment		
Opening applicant's peronal website	0.38	0.33	0.41	-8 (0.03)
At least one piece of information acquired	0.37	0.30	0.40	-10 (0.01)
Number of pieces of information acquired	1.59	1.29	1.75	-0.46 (0.01)
All pieces of information acquired	0.24	0.19	0.26	-8 (0.02)
Likelihood of acquiring information about:				
Education	0.33	0.27	0.36	-9 (0.01)
Habits	0.31	0.26	0.34	-8 (0.01)
Marital status	0.32	0.27	0.35	-8 (0.03)
Job	0.31	0.24	0.35	-11 (0.00)
Age	0.31	0.25	0.34	-9 (0.01)
Obervations	762	258	504	
Panel B				
Sample:		Landlords who opened applicant's website		
At least one piece of information acquired	0.96	0.92	0.98	-6 (0.02)
Number of pieces of information acquired	4.14	3.91	4.24	-0.33 (0.06)
All pieces of information acquired	0.62	0.56	0.64	-7 (0.23)
Likelihood of acquiring information about:				
Education	0.86	0.81	0.88	-6 (0.16)
Habits	0.82	0.78	0.83	-6 (0.27)
Marital status	0.84	0.82	0.85	-3 (0.56)
Job	0.82	0.73	0.86	-13 (0.01)
Age	0.81	0.76	0.83	-6 (0.22)
Obervations	293	85	208	
Panel C				
Sample:		Landlords who opened applicant's website		
Likelihood of opening information about ... first				
Education	0.26	0.21	0.28	-7 (0.24)
Habits	0.20	0.22	0.19	3 (0.55)
Marital status	0.26	0.27	0.25	2 (0.72)
Job	0.15	0.12	0.16	-5 (0.32)
Age	0.09	0.09	0.09	0 (0.94)
Observations	293	85	208	
Panel D				
Sample:		Landlords who acquired all pieces of information		
Order of opening information about ...				
Education	3.06	3.33	2.95	0.38 (0.15)
Habits	3.07	2.85	3.14	-0.29 (0.25)
Marital status	2.60	2.44	2.66	-0.22 (0.33)
Job	2.90	2.88	2.90	-0.03 (0.90)
Age	3.38	3.50	3.34	0.16 (0.47)
Observations	181	48	133	

Notes: Means. Column 4 reports differences in percentage points, in the parentheses we report p-value for a t-test testing the null hypothesis that the difference is zero. The differences in the number of pieces of information acquired on the website and in the order of opening a specific piece of information are reported in absolute terms, not in percentage points.

Table 3.A.6: Czech Rental Housing Market - Responsiveness to Information About Asian and Roma Minority Applicants

Dependent variable: Sample:	Invitation rate	
	Asian minority name (1)	Roma minority name (2)
Monitored Information Treatment	0.09* (0.05)	0.07 (0.05)
Additional text in the email - with high school	0.02 (0.07)	0.12** (0.06)
Additional text in the email - with college	0.19*** (0.07)	0.12** (0.06)
Observations	569	625

Notes: Probit, marginal effects (dF/dx), robust standard errors in parentheses. In both Columns, we control for a dummy variable indicating a landlord being a female, a dummy variable indicating an unknown gender of a landlord (the mean of this variable in the whole sample as well as in the Information with monitoring treatment is 0.02), size of an apartment, price of an apartment rental, and a dummy variable indicating an equipped apartment.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 3.A.7: Czech Rental Housing Market - Education Level and Invitation Rate

Dependent variable	Invitation for an apartment viewing							
	Sample	Treatment with additional text in the email				Monitored Information Treatment, sub-sample of landlords who acquired information about education on applicant's personal webpage		
		All (1)	White majority name (2)	Ethnic minority name (3)	High school degree (4)	College degree (5)	White majority name (7)	Ethnic minority name (8)
Ethnic minority name		-0.30*** (0.06)			-0.29*** (0.06)	-0.22*** (0.06)	-0.15 (0.09)	
College degree		0.01 (0.07)	0.01 (0.06)	0.08 (0.05)			0.18* (0.11)	0.17** (0.08)
Ethnic minority name*College degree		0.07					-0.06	
Constant		0.71*** (0.09)	0.66*** (0.12)	0.45*** (0.11)	0.69*** (0.11)	0.75*** (0.13)	0.89*** (0.13)	0.74*** (0.17)
Observations		587	201	386	311	276	251	69
								182

Notes: OLS in all Columns, standard errors in parentheses. Robust standard errors in parentheses. In all Columns, we control for a dummy variable indicating a landlord being a female, a dummy variable indicating an unknown gender of a landlord (the mean of this variable in the whole sample as well as in the Information with monitoring treatment is 0.02), size of an apartment, price of an apartment rental, and a dummy variable indicating an equipped apartment.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 3.A.8: Czech Rental Housing Market - Education Level and Information Search

	High school degree	College degree	p.p. difference: (p-value)
Panel A			
Sample:	Landlords who opened applicant's personal website and acquired information about education		
Likelihood of acquiring information about ... after information about education is acquired			
Habits	0.78	0.76	2 (0.80)
Marital status	0.86	0.76	10 (0.16)
Job	0.86	0.76	10 (0.15)
Age	0.78	0.77	2 (0.82)
Number of pieces of information acquired after information about education is acquired	4.27	4.1	0.17 (0.55)
Panel B			
Sample:	Landlords who opened personal website of an applicant with White majority name and acquired information about education		
Likelihood of acquiring information about ... after information about education is acquired			
Habits	0.82	0.62	20 (0.18)
Marital status	0.92	0.74	18 (0.23)
Job	0.75	0.63	13 (0.42)
Age	0.77	0.73	4 (0.80)
Number of pieces of information acquired after information about education is acquired	3.83	3.75	0.08 (0.91)
Panel C			
Sample:	Landlords who opened personal website of an applicant with Ethnic minority name and acquired information about education		
Likelihood of acquiring information about ... after information about education is acquired			
Habits	0.77	0.83	-6 (0.45)
Marital status	0.84	0.77	8 (0.34)
Job	0.89	0.83	5 (0.46)
Age	0.79	0.79	0 (0.97)
Number of pieces of information acquired after information about education is acquired	4.35	4.25	0.10 (0.75)

Table 3.A.9: Surveys I and II - Ethnicity and Expected Satisfaction with an Applicant

	W majority name (W) (1)	E minority name (E) (2)	Difference: W-E p-value (3)	Asian minority name (A) (4)	Difference: W-A p-value (5)	Roma minority name (R) (6)	Difference: W-R p-value (7)	Difference: R-A p-value (8)
Panel A: Survey among decision-makers in the rental housing market								
Expected applicant's overall quality	3.57	3.04	0.53 (0.01)	3.04	0.52 (0.03)	3.03	0.53 (0.01)	-0.01 (0.98)
Standard deviation of applicant's expected overall quality	0.63	0.62	0.01 (0.94)	0.62	0.01 (0.94)	0.62	0.01 (0.96)	0.00 (0.99)
Expected informativeness of applicant's personal website	2.66	2.62	0.04 (0.85)	2.55	0.11 (0.63)	2.69	-0.03 (0.88)	0.14 (0.54)
Observations	29	60		31		29		
Panel B: Survey among decision-makers in the labor market								
Expected applicant's overall quality	3.35	2.96	0.39 (0.02)	2.89	0.46 (0.01)	3.02	0.33 (0.10)	0.13 (0.50)
Standard deviation of applicant's expected overall quality	0.55	0.53	0.02 (0.84)	0.49	0.06 (0.63)	0.57	-0.01 (0.91)	0.08 (0.53)
Expected informativeness of applicant's resume	2.97	2.62	0.34 (0.10)	2.62	0.34 (0.11)	2.63	0.34 (0.17)	0.00 (0.99)
Observations	29	61		29		32		

Notes: Means. Panel A reports results of the perception survey among landlords in the rental housing market, Panel B reports results of the perception survey among human resource managers in the labor market. Variable "Expected applicant's overall quality" is measured on a scale 1-5, where 1 means that the decision-maker thinks he/she would be very unsatisfied with the applicant and 5 means very satisfied. The decision-makers were asked to allocate 10 tokens, each representing 10 percent probability, among these five categories of expected overall quality. The variable "Standard deviation of applicant's expected overall quality" is calculated at an individual level, based on allocation of tokens to these five categories. The variable "Expected informativeness of applicant's resume/personal website" is measured on a scale 1-4, where 1 means very uninformative and 4 means very informative. Columns 3, 5 and 7 report differences between applicant's names, in the parentheses we report p-value for a t-test testing the null hypothesis that the difference is zero.

Table 3.A.10: Czech Labor Market - Randomization Check

	White majority name (1)	Pooled Asian and Roma minority name (2)	t-test p-value (3)	Asian minority name (4)	Roma minority name (5)	F-stat p-value (6)
Required high school education	0.90 (0.30)	0.88 (0.33)	0.57	0.89 (0.32)	0.86 (0.35)	0.69
Required previous experience	0.31 (0.47)	0.23 (0.42)	0.13	0.25 (0.44)	0.21 (0.41)	0.26
Sector of sales and services	0.73 (0.44)	0.72 (0.45)	0.74	0.74 (0.44)	0.69 (0.47)	0.73
Application in holiday period	0.23 (0.43)	0.32 (0.47)	0.12	0.31 (0.47)	0.34 (0.48)	0.28
Observations	98	176		99	77	

Notes: Means. Standard deviations in parentheses. Column 3 reports p-value for a t-test testing the null hypothesis that the means are equal for applicants with a majority-sounding name and a minority-sounding name (Asian and Roma minority name). Column 6 reports p-value for an F-test testing the null hypothesis that the means are equal across all three groups of applicants.

Table 3.A.11: Czech Labor Market - Invitation Across Sectors

Dependent variable:	Invitation for a job interview			
	Sales and services		Manual work and administration	
Sample:	(1)	(2)	(3)	(4)
Ethnic minority name	-0.09** (0.05)		-0.14 (0.12)	
Asian minority name		-0.09** (0.04)		-0.12 (0.10)
Roma minority name		-0.05 (0.03)		-0.11 (0.10)
Observations	198	198	51	51

Notes: Probit, marginal effects (dF/dx), robust standard errors in parentheses. In all Columns, we control for dummy variables indicating required high school education, required previous experience, and application being sent during a holiday period (August). In all Columns, the omitted variable is a White majority name.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 3.A.12: Czech Labor Market - Education Level and Invitation Rate

Dependent variable Sample	Invitation for a job interview		
	Employers who opened applicant's resume		
	White majority		Ethnic minority
	All (1)	name (2)	name (3)
Ethnic minority name	-0.10 (0.07)		
College degree	0.01 (0.09)	0.01 (0.10)	0.01 (0.06)
Ethnic minority name*College degree	0.00 (0.11)		
Constant	0.19*** (0.06)	0.19*** (0.07)	0.09** (0.04)
Observations	160	62	98

Notes: OLS in all Columns, standard errors in parentheses. Robust standard errors in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 3.A.13: Czech Labor Market - Information Acquisition

	White majority name (W) (1)	Pooled Asian and Roma minority name (E) (2)	p.p. difference: W-E, (p-value) (3)	Asian minority name (A) (4)	p.p. difference: W-A, (p-value) (5)	Roma minority name (R) (6)	p.p. difference: W-R, (p-value) (7)	p.p. difference: R-A, (p-value) (8)
Panel A								
Sample:				All				
Opening applicant's resume	0.63	0.56	8 (0.22)	0.47	16 (0.03)	0.66	-3 (0.69)	19 (0.01)
Any additional information acquired	0.15	0.13	3 (0.52)	0.11	4 (0.39)	0.14	1 (0.85)	3 (0.53)
Number of pieces of additional information acquired	0.31	0.22	0.08 (0.39)	0.16	0.14 (0.16)	0.3	0.01 (0.96)	0.14 (0.21)
All additional information acquired	0.02	0.01	1 (0.26)	0	2 (0.15)	0.01	1 (0.71)	1 (0.26)
Any additional information acquired (excluding contacts)	0.10	0.06	4 (0.24)	0.04	6 (0.09)	0.09	1 (0.81)	5 (0.17)
Number of pieces of additional information acquired (excluding contacts)	0.19	0.13	0.07 (0.37)	0.07	0.12 (0.13)	0.19	-0.00 (0.99)	0.12 (0.14)
All additional information acquired (excluding contacts)	0.02	0.01	1 (0.26)	0	2 (0.15)	0.01	1 (0.71)	1 (0.26)
Observations	98	176		99		77		
Panel B								
Sample:			Employers who opened applicant's resume					
Any additional information acquired	0.24	0.22	2 (0.80)	0.23	1 (0.92)	0.22	3 (0.74)	-1 (0.83)
Number of pieces of additional information acquired	0.48	0.40	0.09 (0.59)	0.34	0.14 (0.43)	0.45	0.3 (0.87)	0.11 (0.55)
All additional information acquired	0.03	0.01	2 (0.32)	0	3 (0.22)	0.02	1 (0.68)	2 (0.34)
Any additional information acquired (excluding contacts)	0.16	0.11	5 (0.37)	0.09	7 (0.24)	0.14	2 (0.73)	5 (0.42)
Number of pieces of additional information acquired (excluding contacts)	0.31	0.22	0.08 (0.52)	0.15	0.16 (0.28)	0.29	0.01 (0.94)	0.15 (0.32)
All additional information acquired (excluding contacts)	0.03	0.01	2 (0.32)	0	3 (0.22)	0.02	1 (0.68)	2 (0.34)
Likelihood of acquiring information about								
Education	0.08	0.05	3 (0.45)	0.06	2 (0.74)	0.04	4 (0.37)	-2 (0.58)
Job experience	0.13	0.08	5 (0.33)	0.04	9 (0.12)	0.12	1 (0.86)	8 (0.18)
Skills	0.06	0.04	2 (0.51)	0.02	4 (0.29)	0.06	1 (0.90)	4 (0.35)
Hobbies	0.03	0.05	-2 (0.57)	0.02	1 (0.73)	0.08	-5 (0.28)	6 (0.20)
Contacts	0.18	0.17	0 (0.95)	0.19	-1 (0.85)	0.16	2 (0.77)	-3 (0.66)
Qualification	0.16	0.10	6 (0.27)	0.06	10 (0.12)	0.14	2 (0.73)	7 (0.23)
Other characteristics	0.18	0.18	-1 (0.92)	0.19	1 (0.85)	0.18	0 (0.99)	-2 (0.85)
Observations	62	98		47		51		

Notes: Means. Columns 3, 5, 7 and 8 report differences in percentage points, in the parentheses we report p-value for a t-test testing the null hypothesis that the difference is zero. Acquiring more information about qualification is a dummy variable indicating whether an employer clicked on "learn more" buttons on a resume to acquire more information about education, experience, and skills. Acquiring more information about other characteristics is a dummy variable indicating whether she/he acquired more information about hobbies and contact information. The differences in the number of pieces of additional information acquired are reported in absolute terms, not in percentage points.

Table 3.A.14: Survey on Perceptions (Survey III) - Comparison of the Names Used in the Czech Experiments with Other Ethnicity-Signaling Names

Dependent variable	Education level		Quality of housing			
	High school (1)	University (2)	Lodging (3)	Rented flat (4)	Own flat (5)	Own house (6)
Panel A: Majority-sounding names (omitted Jiri Hajek)						
Jan Novotny	0.13 (0.20)	0.01 (0.23)	0.15 (0.21)	-0.02 (0.21)	0.24 (0.20)	0.25 (0.24)
Tomas Svoboda	0.04 (0.20)	0.29 (0.22)	0.17 (0.21)	0.03 (0.21)	0.11 (0.20)	0.27 (0.24)
Jakub Dvorak	0.01 (0.19)	0.13 (0.22)	-0.17 (0.21)	0.07 (0.20)	0.08 (0.20)	0.16 (0.23)
Constant	5.06*** (0.14)	3.71*** (0.16)	1.63*** (0.15)	4.42*** (0.15)	4.00*** (0.14)	3.14*** (0.17)
Observations	324	324	324	324	324	324
Panel B: Asian-sounding names (omitted Phan Quyet Nguyen)						
Pham Hai Xuan	0.14 (0.24)	-0.09 (0.28)	-0.16 (0.26)	0.51** (0.23)	0.28 (0.23)	-0.22 (0.26)
Le Anh Khoi Nguyen	0.05 (0.23)	-0.06 (0.27)	-0.00 (0.26)	0.10 (0.22)	-0.10 (0.23)	-0.30 (0.26)
Hoang Ca Sinh	0.09 (0.24)	0.38 (0.28)	0.14 (0.26)	-0.01 (0.23)	0.03 (0.23)	-0.29 (0.26)
Constant	4.45*** (0.17)	3.31*** (0.19)	2.34*** (0.18)	4.17*** (0.16)	3.47*** (0.16)	2.98*** (0.18)
Observations	330	330	330	329	330	330
Panel C: Roma-sounding names (omitted Gejza Horvath)						
Tibor Farkas	0.94*** (0.27)	0.84*** (0.26)	-0.95*** (0.30)	-0.15 (0.23)	0.48* (0.25)	0.34 (0.23)
Tibor Demeter	0.51* (0.27)	0.63** (0.26)	-0.67** (0.29)	-0.24 (0.22)	-0.09 (0.24)	0.06 (0.23)
Koloman Lakatos	0.16 (0.27)	0.40 (0.27)	-0.18 (0.30)	-0.38 (0.23)	-0.34 (0.25)	0.07 (0.24)
Constant	3.25*** (0.19)	1.63*** (0.19)	4.09*** (0.21)	4.61*** (0.16)	2.63*** (0.17)	1.83*** (0.16)
Observations	322	322	322	320	322	322

Notes: The table reports results of the perception survey about SES among students. OLS in all Columns of all Panels. Standard errors in parentheses. In Panel A, the omitted variable is the name Jiri Hajek, in Panel B it is Phan Quyet Nguyen and in Panel C it is Gejza Horvath. The dependent variables are measured on a scale 0-7. 0 means that a respondent considered it impossible for a person with the given name to have high school (university) education and to live in lodging (in a rented flat, in an own flat, in an own house). 7 means that a respondent considered it certain.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 3.A.15: German Labor Market - Randomization Check

Experimental manipulation:	Name of applicant			Information about unemployment			
	White majority name	Turkish minority name	t-test p-value	No Information	2 months unemployed	18 months unemployed	F-stat p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Required high school education	0.28 (0.45)	0.30 (0.46)	0.67	0.29 (0.45)	0.25 (0.44)	0.33 (0.47)	0.21
Required previous experience	0.57 (0.50)	0.54 (0.50)	0.45	0.53 (0.50)	0.59 (0.49)	0.55 (0.50)	0.39
City with more than 1 million inhabitants	0.18 (0.38)	0.20 (0.40)	0.37	0.18 (0.38)	0.21 (0.41)	0.19 (0.40)	0.61
Application in holiday period	0.17 (0.38)	0.24 (0.43)	0.02	0.21 (0.41)	0.20 (0.40)	0.22 (0.41)	0.92
Sector: manufacturing and construction	0.11 (0.31)	0.09 (0.29)	0.44	0.09 (0.29)	0.10 (0.30)	0.11 (0.32)	0.65
Sector: information and communication	0.17 (0.38)	0.17 (0.37)	0.91	0.19 (0.39)	0.16 (0.36)	0.15 (0.35)	0.42
Sector: administration	0.19 (0.39)	0.15 (0.36)	0.23	0.16 (0.37)	0.20 (0.40)	0.16 (0.36)	0.36
Sector: professional, scientific and technical activities	0.32 (0.47)	0.36 (0.48)	0.20	0.35 (0.48)	0.29 (0.46)	0.37 (0.48)	0.30
Other sector	0.22 (0.42)	0.23 (0.42)	0.79	0.22 (0.41)	0.25 (0.43)	0.22 (0.42)	0.74
Observations	366	379		372	187	186	

Notes: Means. Standard deviations in parentheses. Column 3 reports p-value for a t-test testing the null hypothesis that the means are equal for applicants with a majority-sounding name and a Turkish minority-sounding name. Column 7 reports p-value for an F-test testing the null hypothesis that the means are equal for applicants who do not provide any information about unemployment, for those who say they have been two months unemployed and for those who say they have been a year and a half unemployed.

Bibliography

- Ahmed, Ali M, and Mats Hammarstedt. 2008. "Detecting Discrimination against Homosexuals: Evidence from a Field Experiment on the Internet." *Economica* 76 (303): 588–597.
- Aigner, Dennis J, and Glen G Cain. 1977. "Statistical theories of discrimination in labor markets." *Industrial and Labor Relations Review* 30:175–187.
- Alderman, Harold, and Ruslan Yemtsov. 2014. "How can safety nets contribute to economic growth?" *World Bank Economic Review* 28 (1): 1–20.
- Alesina, Alberto, and Eliana La Ferrara. 2002. "Who trusts others?" *Journal of Public Economics* 85 (2): 207–234.
- . 2005. "Ethnic Diversity and Economic Performance." *Journal of Economic Literature* 43 (3): 762–800.
- Alexander, Marcus, and Fotini Christia. 2011. "Context Modularity of Human Altruism." *Science* 334 (6061): 1392–1394.
- Almås, Ingvild, Alexander W Cappelen, Erik Ø Sørensen, and Bertil Tungodden. 2010. "Fairness and the Development of Inequality Acceptance." *Science* 328 (5982): 1176–1178.
- Altonji, Joseph G, and Rebecca M Blank. 1999. "Race and gender in the labor market." Chapter 48 of *Handbook of Labor Economics*, edited by Orley C Ashenfelter and David Card, Volume 3, 3143–3259. Elsevier.
- Andersen, Steffen, Glenn W Harrison, Morten I Lau, and Elisabet Rutström. 2008. "Lost in state space: Are preferences stable?" *International Economic Review* 49 (3): 1091–1112.
- Andreoni, James, and John Miller. 2002. "Giving according to GARP: An experimental test of the consistency of preferences for altruism." *Econometrica* 70 (2): 737–753.
- Arieli, Amos, Yaniv Ben-Ami, and Ariel Rubinstein. 2011. "Tracking Decision Makers under Uncertainty." *American Economic Journal: Microeconomics* 3 (4): 68–76.
- Arrow, Kenneth. 1973. In *The theory of discrimination*, edited by Orley Ashenfelter and Albert Rees, Volume 3. NJ: Princeton University Press.

- Arrow, Kenneth J. 1972. "Gifts and exchanges." *Philosophy & Public Affairs* 1 (4): 343–362.
- Attanasio, Orazio, Abigail Barr, Juan Camilo Cardenas, Garance Genicot, and Costas Meghir. 2012. "Risk Pooling, Risk Preferences, and Social Networks." *American Economic Journal: Applied Economics* 4 (2): 134–167.
- Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2005. "Social preferences and the response to incentives: Evidence from personnel data." *Quarterly Journal of Economics* 120 (3): 917–962.
- Banerjee, Abhijit, Marianne Bertrand, Saugato Datta, and Sendhil Mullainathan. 2009. "Labor market discrimination in Delhi: Evidence from a field experiment." *Journal of Comparative Economics* 37 (1): 14–27.
- Banerjee, Abhijit V. 2013. "Microcredit Under the Microscope: What Have We Learned in the Past Two Decades, and What Do We Need to Know?" *Annual Review of Economics* 5 (1): 487–519.
- Barany, Zoltan. 2002. *The East European Gypsies: Regime Change, Marginality, and Ethnopolitics*. Cambridge University Press.
- Barr, Abigail. 2003. "Trust and expected trustworthiness: experimental evidence from zimbabwean villages." *Economic Journal* 113 (489): 614–630.
- Barr, Abigail, and Garance Genicot. 2008. "Risk Sharing, Commitment, and Information: An Experimental Analysis." *Journal of the European Economic Association* 6 (6): 1151–1185.
- Basu, Karna, and Maisy Wong. 2015. "Evaluating seasonal food storage and credit programs in east Indonesia." *Journal of Development Economics* 115:200–216.
- Bauer, Michal, Alessandra Cassar, Julie Chytilová, and Joseph Henrich. 2014. "War's Enduring Effects on the Development of Egalitarian Motivations and In-Group Biases." *Psychological Science* 25 (1): 47–57.
- Becker, Garry S. 1971. *The Economics of Discrimination*. Chicago: University of Chicago Press.
- Behrman, Jere R. 1988. "Intrahousehold Allocation of Nutrients in Rural India: Are Boys Favored? Do Parents Exhibit Inequality Aversion?" *Oxford Economic Papers* 40 (1): 32–54.
- Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10 (1): 122–142.
- Bernhard, Helen, Urs Fischbacher, and Ernst Fehr. 2006. "Parochial altruism in humans." *Nature* 442 (7105): 912–915.
- Bertrand, Marianne, Dolly Chugh, and Sendhil Mullainathan. 2005. "Implicit Discrimination." *American Economic Review* 95 (2): 94–98.
- Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94 (4): 991–1013.

- Bhatt, Meghana, and Colin F Camerer. 2005. "Self-referential thinking and equilibrium as states of mind in games: fMRI evidence." *Games and Economic Behavior* 52 (2): 424–459.
- Biggs, Tyler, Mayank Raturi, and Pradeep Srivastava. 2002. "Ethnic networks and access to credit: Evidence from the manufacturing sector in Kenya." *Journal of Economic Behavior & Organization* 49 (4): 473–486.
- Blakeslee, David, and Ram Fishman. 2013. "Rainfall shocks and property crimes in agrarian societies: Evidence from India."
- Block, Jack. 1983. *Lives Through Time*. Berkeley, CA: Bancroft Books.
- Bohnet, Iris, Bruno S Frey, and Steffen Huck. 2001. "More Order with Less Law: On Contract Enforcement, Trust, and Crowding." *American Political Science Review* 95 (01): 131–144.
- Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. "Stereotypes."
- Bornstein, Gary, and Ori Weisel. 2010. "Punishment, cooperation, and cheater detection in "noisy" social exchange." *Games* 1 (1): 18–33.
- Bose, Gautam, and Kevin Lang. 2013. "Monitoring for Worker Quality: Task Assignment, Job Ladders, Wages and Mobility in Nonmanagerial Internal Labor Markets."
- Bowles, Samuel. 2008. "Policies designed for self-interested citizens may undermine "the moral sentiments": Evidence from economic experiments." *Science* 320 (5883): 1605–1609.
- Bowles, Samuel, and Sandra Polania-Reyes. 2012. "Economic incentives and social preferences: Substitutes or complements?" *Journal of Economic Literature* 50 (2): 368–425.
- Boyd, Robert, Herbert Gintis, Samuel Bowles, and Peter J Richerson. 2003. "The evolution of altruistic punishment." *Proceedings of the National Academy of Sciences of the United States of America* 100 (6): 3531–3535.
- Brandts, Jordi, and Gary Charness. 2011. "The strategy versus the direct-response method: a first survey of experimental comparisons." *Experimental Economics* 14:375–398.
- Brocas, Isabelle, Juan D Carrillo, Stephanie W Wang, and Colin F Camerer. 2014. "Imperfect choice or imperfect attention? Understanding strategic thinking in private information games." *Review of Economic Studies* 81 (3): 944–970.
- Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak. 2014. "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh." *Econometrica* 82 (5): 1671–1748.
- Bursell, Moa. 2007. "What's in a Name? A Field Experiment Test for the Existence of Ethnic Discrimination in the Hiring Process."
- Camerer, Colin F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

- Camerer, Colin F, and Eric Johnson. 2004. "Thinking About Attention in Games: Backward and Forward Induction." In *Psychology of Economic Decisions*, edited by Isabel Brocas and Juan Carillo, Volume 2. New York: Oxford University Press.
- Camerer, Colin F, Eric Johnson, Talia Rymon, and Sankar Sen. 1993. "Cognition and framing in sequential bargaining for gains and losses." In *Frontiers of game theory*, edited by Ken G Binmore, Alan P Kirman, and Piero Tani, 27–47. MIT Press.
- Cameron, Lisa, and Manisha Shah. 2015. "Risk-Taking Behavior in the Wake of Natural Disasters." *Journal of Human Resources* 50 (2): 484–515.
- Caplin, Andrew, and Mark Dean. 2015. "Revealed Preference, Rational Inattention, and Costly Information Acquisition." *American Economic Review* 105 (7): 2183–2203.
- Charness, Gary, Luca Rigotti, and Aldo Rustichini. 2007. "Individual Behavior and Group Membership." *American Economic Review* 97 (4): 1340–1352.
- Chetty, Raj, Adam Looney, and Kory Kroft. 2009. "Salience and Taxation: Theory and Evidence." *American Economic Review* 99 (4): 1145–1177.
- Choi, Jung-Kyoo, and Samuel Bowles. 2007. "The Coevolution of Parochial Altruism and War." *Science* 318 (5850): 636–640.
- Coate, Stephen, and Glenn C Loury. 1993. "Will affirmative-action policies eliminate negative stereotypes." *American Economic Review* 83 (5): 1220–1240.
- Coate, Stephen, and Martin Ravallion. 1993. "Reciprocity without commitment." *Journal of Development Economics* 40 (1): 1–24.
- Cohen, Sheldon, Tom Kamarck, and Robin Mermelstein. 1983. "A Global Measure of Perceived Stress." *Journal of Health and Social Behavior* 24 (4): 385–396.
- Collier, Paul. 1999. "The Political Economy of Ethnicity." In *Annual World Bank Conference on Development Economics 1998*, edited by Boris Pleskovic and Joseph E. Stiglitz, 387–399. Washington, D.C.: World Bank Publications.
- Collins, Daryl, Jonathan Morduch, Stuart Rutherford, and Orlanda Ruthven. 2009. *Portfolios of the poor: how the world's poor live on \$2 a day*. Princeton, NJ: Princeton University Press.
- Commission, European. 2010. The Social and Economic Integration of the Roma in Europe.
- Cornell, Bradford, and Ivo Welch. 1996. "Culture, Information, and Screening Discrimination." *Journal of Political Economy* 104 (3): 542–571.
- Costa, Dora L, and Matthew E Kahn. 2003. "Civic Engagement and Community Heterogeneity: An Economist's Perspective." *Perspectives on Politics* 1 (01): 103–111.
- Costa-Gomes, Miguel, Vincent P Crawford, and Bruno Broseta. 2001. "Cognition and Behavior in Normal-Form Games: An Experimental Study." *Econometrica* 69 (5): 1193–1235.
- Costa-Gomes, Miguel A, and Vincent P Crawford. 2006. "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study." *American Economic Review* 96 (5): 1737–1768.

- Cox, James C. 2004. "How to identify trust and reciprocity." *Games and Economic Behavior* 46 (2): 260–281.
- Crawford, Vincent P. 2008. In *Look-ups as the windows of the strategic soul: Studying cognition via information search in game experiments*, edited by Andrew Caplin and Andrew Schotter. Oxford University Press.
- Cronk, Lee, Napoleon A. Chagnon, and William Irons, eds. 2000. *Adaptation and human behavior: an anthropological perspective*. Piscataway, NJ: Transaction Publishers.
- Dechief, Diane, and Philip Oreopoulos. 2012. "Why do some employers prefer to interview Matthew but not Samir? New evidence from Toronto, Montreal and Vancouver."
- Deci, Edward L., and Richard M. Ryan. 1985. *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum Press.
- DellaVigna, Stefano, John A. List, and Ulrike Malmendier. 2012. "Testing for altruism and social pressure in charitable giving." *Quarterly Journal of Economics* 127 (1): 1–56.
- Deresiewicz, William. 2014. *Excellent Sheep: The Miseducation of the American Elite and the Way to a Meaningful Life*. New York, New York: Free Press.
- Devereux, Stephen, Samuel Hauenstein Swan, and Bapu Vaitla. 2008. *Seasons of Hunger: Fighting Cycles of Starvations Among the World's Rural Poor*. London, UK: Pluto Press.
- Devetag, Giovanna, Sibilla Di Guida, and Luca Polonio. 2016. "An eye-tracking study of feature-based choice in one-shot games." *Experimental Economics* 19 (1): 177–201.
- Diamond, Peter A. 1971. "A Model of Price Adjustment." *Journal of Economic Theory* 3:156–168.
- Dirks, Robert. 1980. "Social Responses During Severe Food Shortages and Famine." *Current Anthropology* 21 (1): 21–32.
- Dupas, Pascaline, and Jonathan Robinson. 2013. "Why don't the poor save more? Evidence from health savings experiments." *American Economic Review* 103 (4): 1138–1171.
- Elinder, Mikael, and Oscar Erixson. 2012. "Gender, social norms, and survival in maritime disasters." *Proceedings of the National Academy of Sciences of the United States of America* 109 (33): 13220–13224.
- Engel, Christoph. 2011. "Dictator games: A meta study." *Experimental Economics* 14 (4): 583–610.
- Evans-Pritchard, Edward E. 1969. *The Nuer: A Description of the Modes of Livelihood and Political Institutions of a Nilotic People*. Oxford, UK: Oxford University Press.
- Fafchamps, Marcel. 2000. "Ethnicity and credit in African manufacturing." *Journal of Development Economics* 61 (1): 205–235.
- Falk, Armin, and Urs Fischbacher. 2006. "A theory of reciprocity." *Games and Economic Behavior* 54 (2): 293–315.

- Falk, Armin, and Michael Kosfeld. 2006. "The Hidden Costs of Control." *American Economic Review* 96 (5): 1611–1630.
- Falk, Armin, and Christian Zehnder. 2013. "A city-wide experiment on trust discrimination." *Journal of Public Economics* 100:15–27.
- FAO. 2012. "FAO statistical yearbook 2012: World food and agriculture." Technical Report, Food and Agriculture Organization of the United Nations, Rome, Italy.
- Fehr, E, and S Gächter. 2000. "Cooperation and Punishment in Public Goods Experiments." *American Economic Review* 90 (4): 980–994.
- Fehr, Ernst, Helen Bernhard, and Bettina Rockenbach. 2008. "Egalitarianism in young children." *Nature* 454 (7208): 1079–1083.
- Fehr, Ernst, and Urs Fischbacher. 2003. "The nature of human altruism." *Nature* 425 (6960): 785–791.
- . 2004a. "Social norms and human cooperation." *Trends in Cognitive Sciences* 8 (4): 185–190.
- . 2004b. "Third-party punishment and social norms." *Evolution and Human Behavior* 25 (2): 63–87.
- Fehr, Ernst, and Simon Gächter. 2002. "Altruistic punishment in humans." *Nature* 415 (6868): 137–140 (jan).
- Fehr, Ernst, Georg Kirchsteiger, and Arno Riedl. 1993. "Does Fairness Prevent Market Clearing? An Experimental Investigation." *Quarterly Journal of Economics* 108 (2): 437–459.
- Fehr, Ernst, and John A List. 2004. "The Hidden Costs and Returns of Incentives - Trust and Trustworthiness among CEOs." *Journal of the European Economic Association* 2 (5): 743–771.
- Fehr, Ernst, and Antonio Rangel. 2011. "Neuroeconomic Foundations of Economic Choice - Recent Advances." *Journal of Economic Perspectives* 25 (4): 3–30.
- Fehr, Ernst, and Bettina Rockenbach. 2003. "Detrimental Effects of Sanctions on Human Altruism." *Nature* 422 (6928): 137–140.
- Fehr, Ernst, and Klaus M Schmidt. 1999. "A theory of fairness, competition, and cooperation." *Quarterly Journal of Economics* 114 (3): 817 – 868.
- Fershtman, Chaim, and Uri Gneezy. 2001. "Discrimination in a Segmented Society: An Experimental Approach." *Quarterly Journal of Economics* 116 (1): 351–377.
- Fessler, Daniel M T, and C. David Navarrete. 2004. "Third-party attitudes toward sibling incest: Evidence for Westermarck's hypotheses." *Evolution and Human Behavior* 25 (5): 277–294.
- Fetzer, Thiemo. 2014. "Social Insurance and Conflict: Evidence from India." *Mimeo*.
- Field, Erica. 2007. "Entitled to work: Urban property rights and labor supply in Peru." *Quarterly Journal of Economics* 122 (4): 1561–1602.

- Fiske, Susan T, Amy J C Cuddy, Peter Glick, and Jun Xu. 2002. "A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition." *Journal of Personality and Social Psychology* 82 (6): 878–902.
- Fisman, Raymond, Pamela Jakiela, and Shachar Kariv. 2015. "How did distributional preferences change during the Great Recession?" *Journal of Public Economics* 128:84–95.
- Fisman, Raymond, Pamela Jakiela, Shachar Kariv, and Daniel Markovits. 2015. "The distributional preferences of an elite." *Science* 349 (6254): aab0096–aab0096.
- FRA & UNDP. 2012. The Situation of Roma in 11 EU Member States: Survey Results at a Glance.
- Frey, Bruno S., and Reto Jegen. 2001. "Motivation Crowding Theory." *Journal of Economic Surveys* 15 (5): 589–611.
- Frey, Bruno S., and Felix Oberholzer-Gee. 1997. "The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out." *American Economic Review* 87 (4): 746–755.
- Gabaix, Xavier, David Laibson, and Guillermo Moloche. 2006. "Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model." *American Economic Review* 96 (4): 1043–1068.
- Gintis, Herbert. 2000. "Strong Reciprocity and Human Sociality." *Journal of Theoretical Biology* 206:169–179.
- Gneezy, Ayelet, and Daniel M T Fessler. 2012. "Conflict, sticks and carrots: war increases prosocial punishments and rewards." *Proceedings. Biological sciences / The Royal Society* 279 (1727): 219–223.
- Gneezy, Uri, John List, and Michael K Price. 2012. "Toward an Understanding of Why People Discriminate: Evidence from a Series of Natural Field Experiments."
- Goette, Lorenz, David Huffman, and Stephan Meier. 2012. "The Impact of Social Ties on Group Interactions: Evidence from Minimal Groups and Randomly Assigned Real Groups." *American Economic Journal: Microeconomics* 4 (1): 101–115.
- Goldin, Claudia, and Cecilia Rouse. 2000. "Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians." *American Economic Review* 90 (4): 715–741.
- Grechenig, Kristoffel, Andreas Nicklisch, and Christian Thöni. 2010. "Punishment despite reasonable doubt - a public goods experiment with sanctions under uncertainty." *Journal of Empirical Legal Studies* 7 (4): 847–867.
- Greenwald, Anthony G, Debbie E McGhee, and Jordan L K Schwartz. 1998. "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test." *Journal of Personality and Social Psychology* 74 (6): 1464–1480.
- Greif, Avner. 1993. "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition." *American Economic Review* 83 (3): 525–548.

- Grossman, Herschel I., and Juan Mendoza. 2003. "Scarcity and appropriative competition." *European Journal of Political Economy* 19 (4): 747–758.
- Gürerk, Ozgür, Bernd Irlenbusch, and Bettina Rockenbach. 2006. "The competitive advantage of sanctioning institutions." *Science* 312 (5770): 108–111.
- Habyarimana, James, Macartan Humphreys, Daniel N Posner, and Jeremy M Weinstein. 2007. "Why Does Ethnic Diversity Undermine Public Goods Provision?" *American Political Science Review* 101 (04): 709–725.
- Hare, Todd A, Jonathan Malmaud, and Antonio Rangel. 2011. "Focusing Attention on the Health Aspects of Foods Changes Value Signals in vmPFC and Improves Dietary Choice." *The Journal of Neuroscience* 31 (30): 11077–11087.
- Harrison, Glenn W, and John A List. 2004. "Field Experiments." *Journal of Economic Literature* 42 (4): 1009–1055.
- Haushofer, Johannes, and Ernst Fehr. 2014. "On the psychology of poverty." *Science* 344 (6186): 862–7.
- Heap, Shaun P Hargreaves, and Daniel John Zizzo. 2009. "The value of groups." *American Economic Review* 99 (1): 295–323.
- Henrich, Joseph, and Richard Boyd. 2001. "Why people punish defectors. Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas." *Journal of Theoretical Biology* 208:79–89.
- Henrich, Joseph, Jean Ensminger, Richard McElreath, Abigail Barr, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, Michael Gurven, Edwins Gwako, Natalie Henrich, Carolyn Lesorogol, and Frank an Marlowe. 2010. "Markets, religion, community size, and the evolution of fairness and punishment." *Science* 327 (5972): 1480–4.
- Henrich, Joseph, Richard McElreath, Abigail Barr, Jean Ensminger, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, Michael Gurven, Edwins Gwako, Natalie Henrich, Carolyn Lesorogol, and Frank an Marlowe. 2006. "Costly Punishment Across Human Societies." *Science* 312 (5781): 1767–70.
- Herold, Florian. 2010. "Contractual incompleteness as a signal of trust." *Games and Economic Behavior* 68 (1): 180–191.
- Herrmann, Benedikt, Christian Thöni, and Simon Gächter. 2008. "Antisocial punishment across societies." *Science* 319 (5868): 1362–1367.
- Hidalgo, F Daniel, Suresh Naidu, Simeon Nichter, and Neal Richardson. 2010. "Economic determinants of land invasions." *Review of Economics and Statistics* 92 (3): 505–523.
- Houser, Daniel, Erte Xiao, Kevin McCabe, and Vernon Smith. 2008. "When Punishment Fails: Research on Sanctions, Intentions and Non-cooperation." *Games and Economic Behavior* 62 (2): 509–532.
- Hsiang, Solomon M, Marshall Burke, and Edward Miguel. 2013. "Quantifying the Influence of Climate on Human Conflict." *Science* 341 (6151): 1235367–1–1235367–14.

- Jacquemet, Nicolas, and Constantine Yannelis. 2012. "Indiscriminate discrimination: A correspondence test for ethnic homophily in the Chicago labor market." *Labour Economics* 19:824–832.
- Jalan, Jyotsna, and Martin Ravallion. 1999. "Are the poor less well insured? Evidence on vulnerability to income risk in rural China." *Journal of Development Economics* 58 (1): 61–81.
- Jensen, Robert. 2010. "The (perceived) returns to education and the demand for schooling." *Quarterly Journal of Economics* 125 (2): 515–548.
- Jessor, Richard. 1983. "The Stability of Change: Psychosocial Development From Adolescence to Young Adulthood." In *Human Development: An Interactional Perspective*, edited by David Magnusson and Vernon L. Allen. New York, NY: Academic Press.
- Johansson-Stenman, Olof, Minhaj Mahmud, and Peter Martinsson. 2009. "Trust and religion: Experimental evidence from Rural Bangladesh." *Economica* 76 (303): 462–485.
- Johnson, Noel D, and Alexandra A Mislin. 2011. "Trust games: A meta-analysis." *Journal of Economic Psychology* 32 (5): 865–889.
- Kaas, Leo, and Christian Manger. 2012. "Ethnic Discrimination in Germany's Labour Market: A Field Experiment." *German Economic Review* 13 (1): 1–20.
- Kahneman, Daniel. 1973. *Attention and effort*. Englewood Cliffs, New Jersey: Prentice-Hall.
- . 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.
- Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1986. "Fairness and the Assumptions of Economics." *The Journal of Business* 59 (S4): S285.
- Khandker, Shahidur R, and Wahiduddin Mahmud. 2012. *Seasonal hunger and public policies: Evidence from Northwest Bangladesh*. Washington, DC: World Bank Publications.
- Knack, Stephen, and Philip Keefer. 1997. "Does social capital have an economic payoff? A cross-country investigation." *Quarterly Journal of Economics* 112 (4): 1251–1288.
- Knoepfle, Daniel T, Joseph Tao-yi Wang, and Colin F Camerer. 2009. "Studying learning in games using eye-tracking." *Journal of the European Economic Association* 7 (2-3): 388–398.
- Kocher, Martin G., Peter Martinsson, and Martine Visser. 2008. "Does stake size matter for cooperation and punishment?" *Economics Letters* 99 (3): 508–511.
- Krajibich, Ian, Carrie Armel, and Antonio Rangel. 2010. "Visual fixations and the computation and comparison of value in simple choice." *Nature Neuroscience* 13 (10): 1292–1298.
- Lang, Kevin. 1986. "A Language Theory of Discrimination." *Quarterly Journal of Economics* 101 (2): 363–382.
- Lang, Kevin, and Jee-Yeon K Lehmann. 2012. "Racial Discrimination in the Labor Market: Theory and Empirics." *Journal of Economic Literature* 50 (4): 959–1006.

- Lanjouw, Jean O., and Philip I. Levy. 2002. "Untitled: A study of formal and informal property rights in urban Ecuador." *Economic Journal* 112 (482): 986–1019.
- Leider, Stephen, Markus M Mobius, Tanya Rosenblat, and Quoc-Anh Do. 2009. "Directed Altruism and Enforced Reciprocity in Social Networks." *Quarterly Journal of Economics* 124 (4): 1815–1851.
- Lévesque, Carole, Dominique de Juriew, Catherine Lussier, and Nadine Trudeau. 2000. "Between abundance and scarcity: Food and the institution of sharing among the Inuit of the Circumpolar region during the recent historical period." Chapter 6 of *Sustainable Development in the Arctic Conditions of Food Security*, edited by Duhaime; Gerard, 103–115. Montreal.
- Levitt, Steven D, and John A List. 2007. "What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World ?" *Journal of Economic Perspectives* 21 (2): 153–174.
- Ligon, Ethan, and Laura Schechter. 2012. "Motives for sharing in social networks." *Journal of Development Economics* 99 (1): 13–26.
- List, John A. 2004. "The Nature and Extent of Discrimination in the Marketplace: Evidence from the Field." *Quarterly Journal of Economics* 119 (1): 49–89.
- List, John A, and Imran Rasul. 2011. "Field Experiments in Labor Economics." *Handbook of Labor Economics* 4 (A): 103–228.
- Mackowiak, Bartosz, and Mirko Wiederholt. 2009. "Optimal Sticky Prices under Rational Inattention." *American Economic Review* 99 (3): 769–803.
- Maier-Rigaud, Frank P., Peter Martinsson, and Gianandrea Staffiero. 2010. "Ostracism and the provision of a public good: experimental evidence." *Journal of Economic Behavior & Organization* 73 (3): 387–395.
- Maldonado, Jorge Higinio, Rocío del Pilar Moreno-Sánchez, and Rocio del Pilar. 2009. "Does Scarcity Exacerbate the Tragedy of the Commons?: Evidence from Fishers' Experimental Responses."
- Mani, Anandi, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao. 2013. "Poverty Impedes Cognitive Function." *Science* 341 (6149): 976–980.
- Matějka, Filip, and Alisdair McKay. 2015. "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model." *American Economic Review* 105 (1): 272–298.
- Matějka, Filip, and Christopher A Sims. 2011. "Discrete Actions in Information-Constrained Tracking Problems."
- Meier, Stephan, and Charles D. Sprenger. 2015. "Temporal stability of time preferences." *Review of Economics and Statistics* 97 (2): 273–286.
- Miguel, Edward. 2005. "Poverty and Witch Killing." *Review of Economic Studies* 72 (4): 1153–1172.
- Miguel, Edward, and Mary Kay Gugerty. 2005. "Ethnic diversity, social sanctions, and public goods in Kenya." *Journal of Public Economics* 89 (11-12): 2325–2368.

- Milkman, Katherine L, Modupe Akinola, and Dolly Chugh. 2012. "Temporal Distance and Discrimination: An Audit Study in Academia." *Psychological Science* 23 (7): 710–717.
- Morduch, Jonathan. 1995. "Income Smoothing and Consumption Smoothing." *Journal of Economic Perspectives* 9 (3): 103–114.
- . 2006. "Micro-insurance: The next revolution?" In *What Have We Learned About Poverty?*, edited by Abhijit Banerjee, Roland Benabou, and Dilip Mookherjee, 337–356. Oxford University Press.
- Neumark, David, Roy J Bank, and Kyle D van Nort. 1996. "Sex discrimination in restaurant hiring: An audit study." *Quarterly Journal of Economics* 111 (3): 915–941.
- Newell, Allen, John Calman Shaw, and Herbert A Simon. 1958. "Elements of a theory of human problem solving." *Psychological Review* 65 (3): 151–166.
- Nieuwerburgh, Stijn Van, and Laura Veldkamp. 2010. "Information Acquisition and Under-Diversification." *Review of Economic Studies* 77 (2): 779–805.
- Nikiforakis, Nikos. 2008. "Punishment and counter-punishment in public good games: Can we really govern ourselves?" *Journal of Public Economics* 92:91–112.
- NRVA. 2008. "National Risk and Vulnerability Assessment 2007/8." Technical Report, Ministry of Rehabilitation and Rural Development and the Central Statistics Organisation (CSO), Kabul, Afghanistan.
- Osés-Eraso, Nuria, and Montserrat Viladrich-Grau. 2007. "Appropriation and concern for resource scarcity in the commons: An experimental study." *Ecological Economics* 63 (2-3): 435–445.
- Oster, Emily. 2004. "Witchcraft, Weather and Economic Growth in Renaissance Europe." *Journal of Economic Perspectives* 18 (1): 215–228.
- Ostrom, Elinor, Joanna Burger, Christopher B. Field, Richard B. Norgaard, and David Policansky. 1999. "Revisiting the Commons: Local Lessons, Global Challenges." *Science* 284 (5412): 278–282.
- Pager, Devah. 2007. "The Use of Field Experiments for Studies of Employment Discrimination: Contributions, Critiques, and Directions for the Future." *The Annals of the American Academy of Political and Social Science* 609:104–133.
- Phelps, Edmund S. 1972. "The Statistical Theory of Racism and Sexism." *American Economic Review* 62 (4): 659–661.
- Prediger, Sebastian, Björn Vollan, and Benedikt Herrmann. 2014. "Scarcity and Anti-social Behavior." *Journal of Public Economics* 119:1–9.
- Rabin, Matthew. 1993. "Incorporating fairness into game theory and economics." *American Economic Review* 83 (5): 1281–1302.
- Rashid, Ahmed. 2001. *Taliban - Militant Islam, Oil & Fundamentalism in Central Asia*. New Haven, CT: Yale University Press.
- Ravallion, Martin. 1997. "Famines and economics." *Journal of Economic Literature* 35 (3): 1205–1242.

- Reutskaja, Elena, Rosemarie Nagel, Colin F Camerer, and Antonio Rangel. 2011. "Search Dynamics in Consumer Choice under Time Pressure: An Eye-Tracking Study." *American Economic Review* 101 (2): 900–926.
- Riach, Peter A, and Judy Rich. 2002. "Field Experiments of Discrimination in the Market Place." *Economic Journal* 112 (483): 480–518.
- Sahn, David E. 1989. *Seasonal Variability in Third World Agriculture*. Baltimore, MD: The International Food Policy Research Institute.
- Sekhri, Sheetal, and Adam Storeygard. 2014. "Dowry deaths: Response to weather variability in India." *Journal of Development Economics* 111:212–223.
- Shah, Anwar. 2006. *Local governance in Developing Countries*. Washington, D.C.: The World Bank.
- Sims, Christopher A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics* 50 (3): 665–690.
- Skans, Oskar Nordström, and Olof Åslund. 2012. "Do Anonymous Job Application Procedures Level the Playing Field?" *Industrial and Labor Relations Review* 65 (1): 82–107.
- Spaan, Ernst, Felicitas Hillmann, and Ton van Naerssen. 2005. *Asian Migrants and European Labour Markets: Patterns and Processes of Immigrant Labour Market Insertion in Europe*. Taylor & Francis.
- Stanley, Damian, Elizabeth A Phelps, and Mahzarin Banaji. 2008. "The Neural Basis of Implicit Attitudes." *Current Directions in Psychological Science* 17:164–170.
- Stevens, Mitchell L. 2009. *Creating a Class: College Admissions and the Education of Elites*. Cambridge, Massachusetts: Harvard University Press.
- Tajfel, Henri, M G Billig, R P Bundy, and Claude Flament. 1971. "Social categorization and intergroup behaviour." *European Journal of Social Psychology* 1 (2): 149–178.
- The Economist. 2012. "Ask the expert: How to write a CV." *The Economist*, sep.
- TheLadders. 2012. "Eye Tracking Online Metacognition: Cognitive Complexity and Recruiter Decision Making." Technical Report.
- Tirole, Jean. 2011. "Incomplete Contracts: Where Do We Stand." *Econometrica* 67 (4): 741–781.
- Titmuss, Richard M. 1971. *The Gift Relationship: From Blood Donations to Social Policy*. New York, NY: Pantheon Books.
- Townsend, Robert M. 1994. "Risk and insurance in village India." *Econometrica* 62 (3): 539–591.
- Turnbull, Colin M. 1972. *The Mountain People*. New York, NY: Touchstone.
- Varghese, Shalet Korattukudy, Prakashan Chellattan Veettil, Stijn Speelman, Jeroen Buyse, and Guido Van Huylenbroeck. 2013. "Estimating the causal effect of water scarcity on the groundwater use efficiency of rice farming in South India." *Ecological Economics* 86:55–64.

- Volk, Stefan, Christian Thöni, and Winfried Ruigrok. 2012. "Temporal stability and psychological foundations of cooperation preferences." *Journal of Economic Behavior & Organization* 81 (2): 664–676.
- Voors, Maarten J., Eleonora E. M. Nillesen, Philip Verwimp, Erwin H. Bulte, Robert Lensink, and Daan P. Van Soest. 2012. "Violent Conflict and Behavior: A Field Experiment in Burundi." *American Economic Review* 102 (2): 941–64.
- Weichselbaumer, Doris. 2003. "Sexual orientation discrimination in hiring." *Labour Economics* 10 (6): 629–642.
- Woodford, Michael. 2009. "Information-constrained state-dependent pricing." *Journal of Monetary Economics* 56 (15): S100–S124.
- Wutich, Amber. 2009. "Water Scarcity and the Sustainability of a Common Pool Resource Institution in the Urban Andes." *Human Ecology* 37 (2): 179–192.
- Xiao, Erte, and Howard Kunreuther. 2015. "Punishment and cooperation in stochastic social dilemmas." *Journal of Conflict Resolution*.
- Yinger, John. 1998. "Evidence on Discrimination in Consumer Markets." *Journal of Economic Perspectives* 12 (2): 23–40.